

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 10:57:59 ; Search time 56.312 Seconds
(without alignments)
477.991 Million cell updates/sec

Title: US-10-104-966-12_COPY_779_980
Perfect score: 1050

Sequence: 1 VDDPRSEDRFESHIECRK.....KINLADRMGLSGVQEIKEQ 202

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_101002:*

1:	/SID2/gcgdata/genseq/genseqp-emb1/AA1980.DAT *
2:	/SID2/gcgdata/genseq/genseqp-emb1/AA1981.DAT *
3:	/SID2/gcgdata/genseq/genseqp-emb1/AA1982.DAT *
4:	/SID2/gcgdata/genseq/genseqp-emb1/AA1983.DAT *
5:	/SID2/gcgdata/genseq/genseqp-emb1/AA1984.DAT *
6:	/SID2/gcgdata/genseq/genseqp-emb1/AA1985.DAT *
7:	/SID2/gcgdata/genseq/genseqp-emb1/AA1986.DAT *
8:	/SID2/gcgdata/genseq/genseqp-emb1/AA1987.DAT *
9:	/SID2/gcgdata/genseq/genseqp-emb1/AA1988.DAT *
10:	/SID2/gcgdata/genseq/genseqp-emb1/AA1989.DAT *
11:	/SID2/gcgdata/genseq/genseqp-emb1/AA1990.DAT *
12:	/SID2/gcgdata/genseq/genseqp-emb1/AA1991.DAT *
13:	/SID2/gcgdata/genseq/genseqp-emb1/AA1992.DAT *
14:	/SID2/gcgdata/genseq/genseqp-emb1/AA1993.DAT *
15:	/SID2/gcgdata/genseq/genseqp-emb1/AA1994.DAT *
16:	/SID2/gcgdata/genseq/genseqp-emb1/AA1995.DAT *
17:	/SID2/gcgdata/genseq/genseqp-emb1/AA1996.DAT *
18:	/SID2/gcgdata/genseq/genseqp-emb1/AA1997.DAT *
19:	/SID2/gcgdata/genseq/genseqp-emb1/AA1998.DAT *
20:	/SID2/gcgdata/genseq/genseqp-emb1/AA1999.DAT *
21:	/SID2/gcgdata/genseq/genseqp-emb1/AA2000.DAT *
22:	/SID2/gcgdata/genseq/genseqp-emb1/AA2001.DAT *
23:	/SID2/gcgdata/genseq/genseqp-emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1059	100.0	2227	23	AAE19899	Hepatitis A virus
2	1056	99.7	2227	11	AAR05697	Attenuated hepatitis A virus
3	1056	99.7	2227	18	AAW34074	Hepatitis A virus
4	1056	99.7	2227	21	AAI18607	Amino acid sequence of viral
5	1056	99.7	2227	21	AAI18609	Amino acid sequence of viral
6	1051	99.2	2227	21	AAI18608	Amino acid sequence of viral
7	1021	96.4	2227	7	AAPE0066	Sequence of viral
8	412	38.9	366	6	AAPE0116	Sequence of hepatitis A virus
9	412	38.9	993	6	AAPE0116	Sequence of hepatitis A virus
10	412	38.9	993	6	AAPE0211	Sequence encoded by hepatitis A virus

11	398	37.6	854	6	AAP50287	Sequence encoded b
12	398	37.6	1077	20	AAW55529	A partial hepaditi
13	398	37.6	1091	14	AAAR2246	Translated from 5'
14	277.5	26.2	839	12	AAAR15659	Capsid region of C
15	129	12.2	25	18	AAW42930	Immunogenic Hepati
16	129	12.2	25	18	AAW42969	Immunogenic Hepati
17	129	12.2	25	22	AAAB69448	Synthetic HAV P2A
18	129	12.2	26	22	AAAB69447	Synthetic HAV P2A
19	107	10.1	21	18	AAW42992	Immunogenic Hepati
20	107	10.1	21	22	AAAB69439	Synthetic HAV P2A
21	106	10.0	20	18	AAW42921	Immunogenic Hepati
22	106	10.0	21	22	AAAB69438	Synthetic HAV VP1
23	104	9.8	20	18	AAW42937	Immunogenic Hepati
24	104	9.8	21	22	AAAB69444	Synthetic HAV P2A
25	103	9.7	20	18	AAW42974	Immunogenic Hepati
26	103	9.7	21	22	AAAB69441	Synthetic HAV P2A
27	102	9.6	20	18	AAW42933	Immunogenic Hepati
28	102	9.6	21	22	AAAB69440	Synthetic HAV P2A
29	101	9.5	20	18	AAW42975	Immunogenic Hepati
30	101	9.5	20	18	AAW42926	Immunogenic Hepati
31	101	9.5	20	18	AAW42938	Immunogenic Hepati
32	101	9.5	21	22	AAAB69442	Synthetic HAV P2A
33	101	9.5	21	22	AAAB69443	Synthetic HAV P2A
34	101	9.5	21	22	AAAB69445	Synthetic HAV P2A
35	96	9.1	20	18	AAW42929	Immunogenic Hepati
36	96	9.1	20	22	AAAB69446	Synthetic HAV P2A
37	92.5	8.7	452	21	AAAB69439	Strawberry alcohol
38	92.5	8.7	452	21	AAAV9666	Human protein sequ
39	86.5	8.2	547	22	AAAB95769	N. meningitidis st
40	86	8.2	707	20	AAW93446	Arabidopsis thalia
41	85.5	8.1	271	21	AAAG8508	Arabidopsis thalia
42	85.5	8.1	300	21	AAAG8507	Arabidopsis thalia
43	85.5	8.1	317	21	AAAG8506	Arabidopsis thalia
44	85	8.0	783	21	AAAG51467	Arabidopsis thalia
45	85	8.0	785	21	AAAG51466	Arabidopsis thalia

ALIGNMENTS

```

RESULT 1
ID AAE19899
XX AAE19899
AC AAE19899;
DT 18-JUN-2002 (first entry) ,
XX
XX Hepatitis A virus (HAV) protein.
DE Hepatitis A virus (HAV)
XX
XX Hepatitis A virus; HAV, infection; virucide; fungicide; antibacterial;
KM cyostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
OS
PN MO200213855-A2.
PD 21-FEB-2002.
XX
XX 15-AUG-2001; 2001MO-IB01808.
PF
XX 17-AUG-2000; 2000US-225767P.
PR 29-AUG-2000; 2000US-229175P.
PR 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPER AB.
PA
XX
XX Salberg M, Hultgren C;
PI
XX
XX WPI; 2002-241837/29.
DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT

```

PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus
 XX
 PS Claim 11; Page 82-87; 120pp; English.
 XX
 CC The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to
 CC a hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is hepatitis A virus (HAV) protein.
 CC
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 1059; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1,76-99;
 Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 VDDPSEEDRRFRESHIECKRPYKELRLEVQKQRLKYAOEELSNEVLPPRKKMGJFSQAK 60
 DB 779 VDDPSEEDRRFRESHIECKRPYKELRLEVQKQRLKYAOEELSNEVLPPRKKMGJFSQAK 838
 QY 61 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 839 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
 QY 121 WTEMKDDKIVSLIEKFTSKYKSKVNFPHGMDLEBIANSKDFPMSETDLCFLHMLN 180
 DB 899 WTEMKDDKIVSLIEKFTSKYKSKVNFPHGMDLEBIANSKDFPMSETDLCFLHMLN 958
 QY 181 PKKINLADRMGLSGVQEIKEQ 202
 DB 959 PKKINLADRMGLSGVQEIKEQ 980
 RESULT 2
 AAR05697
 ID AAR05697 standard; protein; 2227 AA.
 AC AAR05697;
 XX
 DT 15-AUG-1990 (first entry)
 XX
 DE Attenuated hepatitis A virus.
 XX
 Hepatitis A virus; vaccine; attenuated.
 XX
 Hepatitis A virus, strain HM-175.
 XX
 FH Key Location/Qualifiers
 FT 1..23
 FT Region /label=VP4 = 1A
 FT 24..245
 FT Region /label=VP2 = 1B
 FT 246..491
 FT Region /label=VP3 = 1C
 FT 492..791
 FT Region /label=VP1 = 1D
 FT 792..980
 FT Region /label=2A
 FT 981..1087
 FT Region /label=2B
 FT 1088..1422
 FT Region /label=2C
 FT 1423..1496
 FT Region /label=3A
 FT 1497..1519
 FT Region /label=3B = VPg
 FT 1520..1738
 FT Region /label=3C
 FT 1739..2227

FT /label=3D
 XX US4894228-A.
 XX 16-JAN-1990.
 PD
 XX 12-JUL-1988; 88US-0217824.
 PF
 XX 12-JUL-1988; 88US-0217824.
 PR 12-JUL-1988; 88US-0652967.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN.
 PA
 XX Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstein SM,
 PI Daemer RJ, Gust ID;
 PI
 XX WPI; 1990-075557/10.
 DR N-PSDB; AA003512.
 XX
 XX Vaccine against hepatitis A virus infection - comprises novel
 PT attenuated hepatitis A virus strain.
 PS
 CC Claim 1; Fig 1; 18pp; English.
 XX
 CC The attenuated HAV is useful for inducing protective immunity against
 CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
 CC several nucleotide changes distributed throughout the genome, is
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
 CC suitable for use as an HAV vaccine. It is noted that not all the changes
 CC are necessary for attenuation and use as a vaccine.
 CC
 XX Sequence 2227 AA;
 SQ
 Query Match 99.7%; Score 1056; DB 11; Length 2227;
 Best Local Similarity 99.5%; Pred. No. 3,4e-99;
 Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 VDDPSEEDRRFRESHIECKRPYKELRLEVQKQRLKYAOEELSNEVLPPRKKMGJFSQAK 60
 DB 779 VDDPSEEDRRFRESHIECKRPYKELRLEVQKQRLKYAOEELSNEVLPPRKKMGJFSQAK 838
 QY 61 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 839 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
 QY 121 WTEMKDDKIVSLIEKFTSKYKSKVNFPHGMDLEBIANSKDFPMSETDLCFLHMLN 180
 DB 899 WTEMKDDKIVSLIEKFTSKYKSKVNFPHGMDLEBIANSKDFPMSETDLCFLHMLN 958
 QY 181 PKKINLADRMGLSGVQEIKEQ 202
 DB 959 PKKINLADRMGLSGVQEIKEQ 980
 RESULT 3
 AAW34074
 ID AAW34074 standard; Protein; 2227 AA.
 AC AAW34074;
 XX
 DT 27-APR-1998 (first entry)
 XX
 DE Hepatitis A virus HM-175 protein sequence.
 XX
 KM HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
 KM infection; vaccine.
 XX
 XX Hepatitis A virus HM-175.
 XX
 FH Key Location/Qualifiers
 FT 1..23
 FT Protein /label= VP4
 FT 24..245
 FT Protein

FT FT /label= VP2
FT Protein 246..491
FT /label= VP3
FT Protein 492..791
FT /label= VP1
FT Protein 792..980
FT /label= 2A
FT Protein 981..1087
FT /label= 2B
FT Protein 1088..1422
FT /label= 2C
FT Protein 1423..1496
FT /label= 3A
FT Protein 1497..1519
FT /label= 3B
FT Protein 1520..1738
FT /label= 3C
FT Protein 1739..2227
FT /label= 3D

MO974016-A2.

30-OCT-1997.

18-APR-1997; 97WO-US06506.

19-APR-1996; 96US-0015642.

(USSH) US SEC DEPT HEALTH.
(USSH) US DEPT HEALTH & HUMAN SERVICES.

Emerson SU, Purcell RH, Raychaudhuri G;

WPI; 1997-535850/49.

N-PSDB; AAT93023.

Human attenuated HAV genome containing simian HAV 2C gene - useful
as vaccines against HAV infection

Disclosure; Fig 13A-D; 66pp; English.

XX This protein sequence is encoded by the human hepatitis A virus
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (I) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (1); (2) a cell transfected with (1) or the RNA
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (1) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.

SO Sequence 2227 AA;

Query Match 99.7%; Score 1056; DB 18; Length 2227;
Best Local Similarity 99.5%; Pred. No. 3.4e-99;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPSEBDRRPSHIECKRPYKELRLEVGGKORLKYAOBELSNEVLPPPRKKKGJFSQAK 60
Db 779 VDDPSEBDRRPSHIECKRPYKELRLEVGGKORLKYAOBELSNEVLPPPRKKKGJFSQAK 838
Qy 61 ISLFTTEHEIMKFSWGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
Db 839 ISLFTTEHEIMKFSWGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
Qy 121 WTEMKDDKIVSLIEKFTSNKYSKVPFHGMULDLEIANSKDPFNMSETDLCFLHMLN 180

Db 899 WTEMKDDKIVSLIEKFTSNKYSKVPFHGMULDLEIANSKDPFNMSETDLCFLHMLN 958
Qy 181 PKKINLADRMGLSGVQEIKEQ 202
Db 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 4
AAB18607 standard; Protein; 2227 AA.

15-JAN-2001 (first entry)

Amino acid sequence of wild type Hepatitis A virus strain HM-175.

HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

Hepatitis A virus.

US6113912-A.

05-SEP-2000.

07-JUN-1995; 95US-0475886.

18-SEP-1992; 92US-0947338.
17-SEP-1993; 93WO-US08610.
10-MAR-1995; 95US-0397232.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

WPI; 2000-586464/55.
N-PSDB; AAA75476.

Novel live hepatitis A virus adapted to growth in human fibroblast cell
line useful as vaccine for protecting humans against hepatitis A virus
infection, has modified genome compared to wild type

Disclosure; Fig 6A-K; 72pp; English.

CC The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.

SO Sequence 2227 AA;

Query Match 99.7%; Score 1056; DB 21; Length 2227;
Best Local Similarity 99.5%; Pred. No. 3.4e-99;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPSEBDRRPSHIECKRPYKELRLEVGGKORLKYAOBELSNEVLPPPRKKKGJFSQAK 60
Db 779 VDDPSEBDRRPSHIECKRPYKELRLEVGGKORLKYAOBELSNEVLPPPRKKKGJFSQAK 838
Qy 61 ISLFTTEHEIMKFSWGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
Db 839 ISLFTTEHEIMKFSWGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
Qy 121 WTEMKDDKIVSLIEKFTSNKYSKVPFHGMULDLEIANSKDPFNMSETDLCFLHMLN 180
Db 899 WTEMKDDKIVSLIEKFTSNKYSKVPFHGMULDLEIANSKDPFNMSETDLCFLHMLN 958
Qy 181 PKKINLADRMGLSGVQEIKEQ 202
Db 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 5

AAB18609

ID AAB18609 standard; Protein; 2227 AA.

AC AAB18609;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

KM HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection; HAV 4380.

OS Hepatitis A virus.

PN US6113912-A.

PF 05-SEP-2000. 95US-0475886.

PR 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

DR WPI; 2000-586464/55.

DR N-PSDB; AAA75478.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

PT infection, has modified genome compared to wild type

PS Disclosure; Columns 93-104; 72pp; English.

CC The present sequence is derived from a live attenuated hepatitis A

CC virus (HAV) of the invention, designated HAV 4380. The sequence is

CC produced by modifying wild type HAV strain HM-174. The HAV of the

CC invention are adapted to growth in the human fibroblast-like cell

CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain

CC appropriate attenuation. It is useful as a live vaccine for prophylaxis

CC of hepatitis A in humans and other primates.

CC

CC

CC

CC

CC

CC

CC

CC

CC

Sequence 2227 AA;

Query Match 99.7%; Score 1056; DB 21; Length 2227;

Best Local Similarity 99.5%; Pred. No. 3.4e-99;

Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECKRPYKELRLEVQKORLKYAOEELSNEVLPPPRKMKGLFSQAK 60

DB 779 VDDPSEEDRRFRESHIECKRPYKELRLEVQKORLKYAOEELSNEVLPPPRKMKGLFSQAK 838

QY 61 ISLFYTEEHEIMKFSWGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120

DB 839 ISLFYTEEHEIMKFSWGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898

QY 121 WTEMKDDKIVSLIEKFTSNKYWSKYNPPHGMJLDEIANSKDPNMSBTDLCFLLHWIN 180

DB 899 WTEMKDDKIVSLIEKFTSNKYWSKYNPPHGMJLDEIANSKDPNMSBTDLCFLLHWIN 958

QY 181 PKKINLADRMJLGLSGVOEIKEQ 202

DB 959 PKKINLADRMJLGLSGVOEIKEQ 980

RESULT 6

AAB18608

ID AAB18608 standard; Protein; 2227 AA.

AC AAB18608;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

KM HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection; P-35 virus.

OS Hepatitis A virus.

PN US6113912-A.

PF 05-SEP-2000. 95US-0475886.

PR 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

DR WPI; 2000-586464/55.

DR N-PSDB; AAA75477.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

PT infection, has modified genome compared to wild type

PS Disclosure; Columns 67-78; 72pp; English.

CC The present sequence is derived from passage 35 of a wild type

CC hepatitis A virus (HAV) strain HM-174. The resulting virus is

CC designated P-35 virus. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

CC The HAV is able to propagate in MRC-5 cells and retain appropriate

CC attenuation. It is useful as a live vaccine for prophylaxis of

CC hepatitis A in humans and other primates.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Sequence 2227 AA;

Query Match 99.2%; Score 1051; DB 21; Length 2227;

Best Local Similarity 99.0%; Pred. No. 1.1e-98;

Matches 200; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECKRPYKELRLEVQKORLKYAOEELSNEVLPPPRKMKGLFSQAK 60

DB 779 VDDPSEEDRRFRESHIECKRPYKELRLEVQKORLKYAOEELSNEVLPPPRKMKGLFSQAK 838

QY 61 ISLFYTEEHEIMKFSWGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120

DB 839 ISLFYTEEHEIMKFSWGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898

QY 121 WTEMKDDKIVSLIEKFTSNKYWSKYNPPHGMJLDEIANSKDPNMSBTDLCFLLHWIN 180

DB 899 WTEMKDDKIVSLIEKFTSNKYWSKYNPPHGMJLDEIANSKDPNMSBTDLCFLLHWIN 958

QY 181 PKKINLADRMJLGLSGVOEIKEQ 202

DB 959 PKKINLADRMJLGLSGVOEIKEQ 980

RESULT 7

AAB60066

ID AAB60066 standard; Protein; 2227 AA.

AC AAB60066;

DT	26-JUN-1991	(first entry)
DE	Sequence of viral 1434 polypeptide encoded by the complete	
DE	nucleotide sequence of the HAV genome.	
XX		
XX	Diagnosis; vaccine; passive immunotherapy.	
XX		
OS	Hepatitis A virus.	
FT		
FT	Key	Location/Qualifiers
FT	Region	1..245
FT		/label= P1.1A
FT	Region	246..491
FT		/label= 1B
FT	Region	492..836
FT		/label= 1C
FT	Region	837..980
FT		/label= P2.2A
FT	Region	981..1076
FT		/label= 2B
FT	Region	1077..1422
FT		/label= 2C
FT	Region	1423..1484
FT		/label= P3.3A
FT	Region	1485..1507
FT		/label= 3B
FT	Region	1508..1678
FT		/label= 3C
FT	Region	1679..2227
FT		/label= 3D
XX		
PN	EP199480-A.	
XX		
PD	29-OCT-1986.	
XX		
PE	03-APR-1986; 86EP-0302465.	
XX		
PR	03-APR-1985; 85US-0719329.	
XX		
PA	(CHIR-) CHIRON CORP.	
XX		
PI	Dina D, Potter SJ, Vanneest GA, Caput D;	
XX		
DR	WPI; 1986-286213/44.	
DR	N-PSDB; AAN60080.	
XX		
FT	Hepatitis A virus nucleotide sequence and polypeptide - and use	
FT	in prodn. of vaccines and diagnostic probes	
XX		
XX	Claim 5; Fig 1; 18pp; English.	
XX		
CC	AAN60080 and oligonucleotide fragments are useful in detection of	
CC	hepatitis A virus; transformed hosts may be used for expression of	
CC	polypeptides and fragments useful in vaccines without risk of	
CC	infection by the virus or in prodn. of particles which are capable	
CC	of inducing immunocompetent B cells for passive immunotherapy. Pref.	
CC	epitope is derived from Aas 445-657 or 792-848 of the HAV	
CC	polypeptide sequence (AAN60066).	
CC		
XX		
SQ	Sequence 2227 AA;	
Query Match	96.4%; Score 1021; DB 7; length 2227;	
Best local similarity	96.0%; Pred. No. 14e-95;	
Matches 194; Conservative	4; Mismatches 4; Indels 0; Gaps	0.
QY	1 VDDPSEEDRPFESHI ECRKPYKEURLLEVQKORLKYAOEELSENEVLPPRRKMGJFSQAK 60	
Db	779 VDDPSEEDRPFESHI ECRKPYKEURLLEVQKORLKYAOEELSENEVLPPRRKMGJFSQAK 838	
QY	61 ISLFTEHEIMKFSWRGVTADTRALRRRGFSILAAGRSVWTEMDAGVLTGRILINDEK 120	
Db	839 ISLFTEHEIMKFSWRGVTADTRALRRRGFSILAAGRSVWTEMDAGVLTGRILINDEK 898	

Qy	121	WTENMDDDKIVSLIEKFTSNKTKWSKINPFHGMGLDEEIAANSNDPNNSETDLCILLHMLN	180
Db	899	WTENMDDDKIVSLIEKFTSNKTKWSKINPFHGMGLDEEIAANSNDPNNSETDLCILLHMLN	958
Qy	181	PKKINLADRMGLSGVQEIKEQ 202	
Db	959	PKKINLADRMGLSGVQEIKEQ 980	
RESULT 8			
ID	AAP50230	AAP50230 standard; Protein; 366 AA.	
XX			
AC	AAP50230;		
XX			
DT	28-NOV-1991	(first entry)	
XX			
DE		Sequence of hepatitis A virus (HAV) surface protein (VP-1).	
XX			
KW		Hepatitis A virus vaccine; immunisation; monoclonal antibody;	
XX		diagnostic assay.	
OS		Hepatitis A virus.	
XX			
PN	EP138704-A.		
XX			
PD	24-APR-1985.		
XX			
PF	09-OCT-1984;	84EP-0402025.	
XX			
PR	02-MAR-1984;	84US-0585942.	
XX			
PR	14-OCT-1983;	83US-0541836.	
XX			
PA	(MERI)	MERCK & CO INC.	
XX			
PI	Hughes JV, Scolnick EM, Tomassini JE;		
XX			
DR	WPI; 1985-100818/17.		
DR	N-PSDB; AAN50274.		
XX			
PT		New hepatitis A virus surface protein - useful for binding to	
PT		neutralising antibodies to the virus	
XX			
PS	Claim 21; Page 46-48; 49pp;	English.	
XX			
CC		VP1 is isolated by solubilisation of the intact virus in an aq.	
CC		antonic surfactant and a reducing agent. The viral proteins are sepd.	
CC		and the protein of molecular wt. 33000 daltons is sepd.	
XX			
SQ	Sequence	366 AA;	
Query Match			
	Best Local Similarity	38.9%; Score 412; DB 6; Length 366;	
	Matches 78; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1	VDDPSEEDRRFESHIECRKPYKEIRLEVGQRQLTKYAOEISNEVLPPIKWKGLFSQAK 60	
Db	288	VDDPSEEDRRFESHIECRKPYKEIRLEVGQRQLTKYAOEISNEVLPPIKWKGLFSQAK 347	
Qy	61	ISLFYTEHEHIMKFSWRCV 79	
Db	348	ISLFYTEHEHIMKFSWRCV 366	
RESULT 9			
ID	AAP50116	AAP50116 standard; Protein; 993 AA.	
XX			
AC	AAP50116;		
XX			
DT	30-SEP-1991	(first entry)	
XX			
DE		Sequence of Hepatitis A virus (HAV) immunogenic peptides	

DE VP-1, VP-2, VP-3 and VP-4.
 XX Antigenic protein; immunogen; vaccine.
 XX Hepatitis A virus (strain CR326).
 OS
 XX EPI54587-A.
 PN
 XX 11-SEP-1985.
 PD
 XX 27-FEB-1985; 85EP-0400369.
 PF
 XX 02-MAR-1984; 84US-0585818.
 PR
 XX (MERI) MERCK & CO INC.
 PA
 XX Linemeyer DL, Menke JG, Reuben RG, Mitra SW;
 PI
 XX WPI; 1985-224964/37.
 DR
 XX N-PSDB; AAN50139.
 DE New nucleotide sequences coding for hepatitis A virus antigens -
 PI useful for eliciting normal immune response and in vaccines for
 PT protecting against the virus
 PS
 XX Example; Page 11-17; 32pp; English.
 CC Within the sequence in AAN50139 is encoded the information necessary
 CC to make the antigenic proteins of HAV. The sequences encoding for
 CC the structural proteins begin at base 403. The key sub-unit
 CC sequences within VP-1, designated Sequences I, II, III, IV, and V,
 CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
 CC nucleotide sequences which are valuable as encoding antigenic
 CC proteins are the sequences from base 1749 to base 2722, from base
 CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
 CC base 1749 to base 2722 is esp. valuable as a vector for producing
 CC antigen protein. Sequences II-V are claimed. X in AAN50116 denotes the
 CC translation of a stop codon.
 CC
 SQ Sequence 993 AA:
 Query Match 38.9%; Score 412; DB 6; Length 993;
 Best Local Similarity 98.7%; Pred. No. 2.6e-33;
 Matches 78; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VDDPRSEEDRRFESHIECRKPYKELRLVGGKORLKYAOEELSNEVLPPPRKMGGLFSQAK 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 915 VDDPRSEEDRRFESHIECRKPYKELRLVGGKORLKYAOEELSNEVLPPPRKMGGLFSQAK 974
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 ISLFYTEHEHIMKFSWGRV 79
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 975 ISLFYTEHEHIMKFSWGRV 993
 Ub
 RESULT 10
 AAP50231
 ID AAP50231 standard; Protein; 993 AA.
 XX
 AC AAP50231;
 PN
 XX 28-NOV-1991 (first entry)
 PD
 XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).
 XX
 XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.
 KM
 XX Hepatitis A virus.
 OS
 XX
 PI Key Location/Qualifiers
 FH 628..993
 FT Protein /note= "claimed; X denotes translated stop codons
 FT

FT and unspecified triplets"
 XX EPI38704-A.
 PN
 XX 24-APR-1985.
 PD
 XX 09-OCT-1984; 84EP-0402025.
 PF
 XX 02-MAR-1984; 84US-0585942.
 PR
 XX 14-OCT-1983; 83US-0541836.
 PA
 XX (MERI) MERCK & CO INC.
 PI
 XX Hughes JV, Scolnick EM, Tomassini JE;
 DR
 XX WPI; 1985-100818/17.
 DR
 XX N-PSDB; AAN50274.
 DE New hepatitis A virus surface protein - useful for binding to
 PI neutralising antibodies to the virus
 PS
 XX Disclosure; Page 17-23; 49pp; English.
 CC VPI is isolated by solubilisation of the intact virus in an aq.
 CC antionic surfactant and a reducing agent. The viral proteins are sepd.
 CC and the protein of molecular wt. 33000 daltons is sepd.
 CC
 SQ Sequence 993 AA:
 Query Match 38.9%; Score 412; DB 6; Length 993;
 Best Local Similarity 98.7%; Pred. No. 2.6e-33;
 Matches 78; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VDDPRSEEDRRFESHIECRKPYKELRLVGGKORLKYAOEELSNEVLPPPRKMGGLFSQAK 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 915 VDDPRSEEDRRFESHIECRKPYKELRLVGGKORLKYAOEELSNEVLPPPRKMGGLFSQAK 974
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 ISLFYTEHEHIMKFSWGRV 79
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 975 ISLFYTEHEHIMKFSWGRV 993
 Db
 RESULT 11
 AAP50287
 ID AAP50287 standard; Protein; 854 AA.
 XX
 AC AAP50287;
 PN
 XX 30-NOV-1991 (first entry)
 PD
 XX Sequence encoded by hepatitis A virus (HAV) cDNA from near the
 DE genome 5' terminus to the end of the area corresponding to the
 DE capsid protein region of poliovirus RNA.
 DE
 XX Hepatitis A virus assay; antigen; antibody.
 KW
 XX Hepatitis A virus.
 OS
 XX
 PN WO8501517-A.
 PD
 XX 11-APR-1985.
 PF
 XX 27-SEP-1984; 84WO-US01552.
 PR
 XX 30-SEP-1983; 83US-0537911.
 PA
 XX (MASI) MASSACHUSETTS INST TECH.
 PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 FH Racanelli VR;
 DR
 XX WPI; 1985-098846/16.
 DR
 XX N-PSDB; AAN50330.
 DR

```
XX New hepatitis A virus cDNA - useful in assays for the virus and
PT for prodn. of the viral antigen and antibodies to it
XX
PS Example; Fig 7; 60pp; English.
CC The inventors claim HAV cDNA and a method for producing it, whereby
CC large amts. can be obtd. economically. The cDNA is useful in the
CC assay for detection of HAV quickly and easily and with high
CC sensitivity and specificity. The HAV cDNA is also used in the prodn.
CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.
XX
SQ Sequence 854 AA;

Query Match 37.6%; Score 398; DB 6; Length 854;
Best Local Similarity 98.7%; Pred. No. 5,8e-32;
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAQEELSNEVLPPPRKKMGLFSQAK 60
779 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAQEELSNEVLPPPRKKMGLFSQAK 838
DB 61 ISLFYTEHEHMKFSW 76
839 ISLFYTEHEHMKFSW 854

RESULT 12
AAM95559
ID AAM95559 standard; Protein; 1077 AA.
AC AAM95559;
XX
XX 28-APR-1999 (first entry)
DE A partial hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus protein; HAV; P2 region;
KW cell-culture-adapted HAV strain; infection; accelerated growth.
XX
XX Hepatitis A virus.
OS
XX
XX US5849562-A.
PN
XX
XX 15-DEC-1998.
PD
XX
XX 06-JUN-1995; 95US-0468926.
PF
XX
XX 06-NOV-1991; 91US-0788262.
XX 30-SEP-1983; 83US-0537911.
XX 27-SEP-1984; 84US-0654942.
XX 06-OCT-1988; 88US-0256135.
XX 06-JUN-1995; 95US-0468926.
PR
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Emerson SU, Purcell RH;
PI
XX
XX WPI; 1999-094412/08.
XX N-PSDB; AAX01006.
DR
XX
XX Chimeric hepatitis A virus strains - with P2 region from
PT cell-culture-adapted strain in wild-type genome
XX
XX
PS Disclosure; Fig 7A-L; 36pp; English.
XX
XX The present sequence represents a partial hepatitis A virus (HAV)
CC protein. The specification describes a DNA construct consisting
CC of a wild-type HAV genome in which the P2 region is replaced by the
CC P2 region from a cell-culture-adapted HAV strain. The construct is
CC used to demonstrate that mutations in the P2 region of a
CC cell-culture-adapted HAV strain are sufficient for establishment of
CC infection and accelerated growth in cell culture.
```

```
XX
SQ Sequence 1077 AA;

Query Match 37.6%; Score 398; DB 20; Length 1077;
Best Local Similarity 98.7%; Pred. No. 8e-32;
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAQEELSNEVLPPPRKKMGLFSQAK 60
DB 1002 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAQEELSNEVLPPPRKKMGLFSQAK 1061
QY 61 ISLFYTEHEHMKFSW 76
DB 1062 ISLFYTEHEHMKFSW 1077

RESULT 13
AAR32426
ID AAR32426 standard; Protein; 1091 AA.
AC AAR32426;
XX
XX 17-DEC-2001 (updated)
DT 10-JUN-1993 (first entry)
XX
XX Translated from 5' region of Hepatitis A Virus genomic clone.
DE
XX
XX HAV HM-175; chronic liver disease; picornavirus.
XX
XX Hepatitis A virus.
OS
XX
XX Key Location/Qualifiers
FH 238..1091
FT Region /label= ORF
FT /note= "second putative initiation codon at
FT position 240"
FT 1..711
FT /note= "X's correspond to nonsense codons,
FT i.e. this region is not an ORF"
XX
XX USN7788262-N.
PN
XX
XX 15-DEC-1992.
PD
XX
XX 30-SEP-1983; 83US-0536911.
PF
XX
XX 27-SEP-1984; 84US-0654942.
XX 06-OCT-1988; 88US-0256135.
XX 30-SEP-1983; 83US-0536911.
XX 06-NOV-1991; 91US-0788262.
PR
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PA
XX
XX Baltimore D, Feinstone SM;
PI Purcell RH, Racanelli VR, Ticehurst JR;
XX
XX WPI; 1993-067429/08.
XX N-PSDB; AAQ36934.
DR
XX
XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
PT of antigen and antibodies
XX
XX
PS Disclosure; Fig 7; 65pp; English.
XX
XX HAV virion RNA was extracted from the livers of marmosets which had
CC been inoculated with HAV (the HAV had previously been passaged twice
CC in marmosets). The RNA was used to prepare ds cDNA clones by
CC standard methods. Clones concg. inserts which hybridised to RNA from
CC HAV-infected African Green Monkey kidney cells were selected for
CC further analysis. A 7.4kb restriction map (about 99% of the HAV
CC genome) was constructed from 5 overlapping inserts. The sequence of
CC the first 3.3kb (approx.) from the 5'-terminus was determined. An
CC amino acid sequence was deduced from the entire clone and an open
```

CC reading frame was identified starting at position 238. A comparison
 CC of the predicted HAV amino acid sequences with the known capsid
 CC protein sequences of other picornaviruses (poliovirus, foot and
 CC mouth disease virus and encephalomyelitis virus) revealed areas of
 CC local homology.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government-owned NTIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)

XX Sequence 1091 AA;

Query Match 37.6%; Score 398; DB 14; Length 1091;
 Best Local Similarity 98.7%; Pred. No. 8,1e-32;

Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEDRRPFESHIIECKRPYKELRLEVGVKQRLKYAOEISNEVLPPPRKMKGLFSQAK 60

DB 1016 VDDPRSEDRRPFESHIIECKRPYKELRLEVGVKQRLKYAOEISNEVLPPPRKMKGLFSQAK 1075

61 ISLFTTEHEHMKFSW 76

DB 1076 ISLFTTEHEHMKFSW 1091

RESULT 14

AA015629

AA015629 standard; Protein; 839 AA.

AA015629;

17-DEC-2001 (updated)

17-MAR-1992 (first entry)

Capsid region of cyto-HAV isolate CY-145.

Hepatitis A virus; cynomolgus; HAV; monkey; vaccine; macaque.

Cynomolgus monkey hepatitis A virus, isolate CY-145.

Key

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

Protein

XX Disclosure; Fig 3; 23pp; English.

CC The sequence was deduced from the nucleotide sequence obtd. by PCR
 CC amplification of cyto-HAV viral RNA obtd. from the stool of a
 CC cynomolgus monkey with serologically and histologically confirmed
 CC spontaneous hepatitis A. The sequence differs from the human HAV
 CC isolate HM175 (Cohen, J.I., et al. (1987) Proc. Natl. Acad. Sci.
 CC USA 84, 2497-2501), mainly in the VP3 and VP1 proteins. The Gln-Val
 CC pair at the VP3-VP1 cleavage site in the human isolate is replaced
 CC by a Gln-Thr pair in the cyto-HAV. The other two cleavage sites are
 CC the same. Two residues have been identified as part of the immuno-
 CC dominant region (see feature table) and are different to those in
 CC the same position in human HAV. The protein and peptides derived
 CC from it can be used in the prepn. of vaccines for the prevention of
 CC HAV infection.

CC See also AA015056.

CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government-owned NTIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)

XX Sequence 839 AA;

Query Match 26.2%; Score 277.5; DB 12; Length 839;
 Best Local Similarity 85.5%; Pred. No. 1.5e-19;

Matches 53; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 1 VDDPRSEDRRPFESHIIECKRPYKELRLEVGVKQRLKYAOEISNEVLPPPRKMKGLFSQAK 60

DB 779 VDDPRSEDRRPFESHIIECKRPYKELRLEVGVKQRLKYAOEISNEVLPPPRKMKGLFSQAK 837

QY 61 IS 62

DB 838 IS 839

RESULT 15

AA042930

AA042930 standard; peptide; 25 AA.

AA042930;

28-APR-1998 (first entry)

Immunogenic Hepatitis A virus peptide YK-1665.

Immunogenic peptide; immunogenic epitope; P2A protein;
 immune response; antibody.

Synthetic.

Hepatitis A virus.

WO9740147-A1.

30-OCT-1997.

18-APR-1997; 97WO-US06891.

19-APR-1996; 96US-0015644.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Fields HA, Knudskov YE;

WPI; 1997-535831/49.

Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
 immune response to HAV in a mammal or to detect the presence of
 antibodies against HAV in a mammal

Claim 18; Page 112; 140pp; English.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compuser Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 11:05:59 ; Search time 11.8824 Seconds
(without alignments)
500.190 Million cell updates/sec

Title: US-10-104-966-12_COPY_779_980
Perfect score: 1059
Sequence: 1 VDDPRSEDRPFESHIECRK.....KINLADRMGLSGVOEIKEQ 202

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1056	99.7	2227	3 US-08-475-886-2	Sequence 2, Appli
2	1056	99.7	2227	3 US-08-475-886-6	Sequence 6, Appli
3	1056	99.7	2227	4 US-08-397-232-2	Sequence 2, Appli
4	1056	99.7	2227	4 US-08-397-232-4	Sequence 4, Appli
5	1056	99.7	2227	4 US-09-171-387-2	Sequence 2, Appli
6	1056	99.7	2227	4 US-09-653-499-2	Sequence 2, Appli
7	1056	99.7	2227	4 US-09-653-499-6	Sequence 6, Appli
8	1051	99.2	2227	3 US-08-475-886-4	Sequence 4, Appli
9	1051	99.2	2227	4 US-09-653-499-4	Sequence 4, Appli
10	398	37.6	1091	6 5516630-2	Patent No. 5516630
11	277.5	26.2	839	4 US-08-087-016-2	Sequence 2, Appli
12	84.5	8.0	1353	4 US-09-398-193-99	Sequence 99, Appli
13	82.5	7.8	643	2 US-08-245-511-47	Sequence 47, Appli
14	82.5	7.8	1294	2 US-08-600-993A-47	Sequence 47, Appli
15	82.5	7.8	1294	4 US-09-473-717-2	Sequence 4, Appli
16	81.5	7.7	976	4 US-09-104-324B-4	Sequence 4, Appli
17	81	7.6	229	4 US-09-248-528-17	Sequence 17, Appli
18	81	7.6	229	4 US-09-549-108-17	Sequence 17, Appli
19	81	7.6	229	4 US-09-549-111-17	Sequence 17, Appli
20	81	7.6	229	4 US-09-549-106-17	Sequence 17, Appli
21	81	7.6	229	4 US-09-550-394-17	Sequence 17, Appli
22	81	7.6	582	4 US-08-194-560-2	Sequence 2, Appli
23	79.5	7.5	1250	4 US-08-441-139-9	Sequence 9, Appli
24	79.5	7.5	1305	4 US-08-864-785-3	Sequence 2, Appli
25	79.5	7.5	1353	3 US-08-894-173-2	Sequence 2, Appli
26	79.5	7.5	1353	4 US-09-398-193-2	Sequence 2, Appli
27	79.5	7.5	1353	4 US-09-473-717-3	Sequence 3, Appli

28	77.5	7.3	705	4 US-09-370-838-186	Sequence 186, App
29	77.5	7.3	1440	4 US-09-357-251-37	Sequence 37, Appli
30	77.5	7.3	1512	4 US-09-443-184-48	Sequence 48, Appli
31	76.5	7.2	370	2 US-08-837-593-7	Sequence 7, Appli
32	76	7.2	3443	2 US-08-416-603-2	Sequence 2, Appli
33	74	7.0	400	4 US-08-961-083-190	Sequence 190, App
34	73.5	6.9	513	4 US-09-724-519-6	Sequence 6, Appli
35	73.5	6.9	513	4 US-09-592-037-6	Sequence 6, Appli
36	73.5	6.9	575	4 US-09-724-519-8	Sequence 8, Appli
37	73.5	6.9	575	4 US-09-592-037-8	Sequence 8, Appli
38	73.5	6.9	1001	4 US-09-060-410-2	Sequence 2, Appli
39	73	6.9	481	4 US-08-928-862-3	Sequence 3, Appli
40	73	6.9	600	4 US-09-423-468A-13	Sequence 13, Appli
41	72	6.8	877	4 US-09-165-386-5	Sequence 5, Appli
42	72	6.8	1259	4 US-09-134-001C-3757	Sequence 3757, Ap
43	70.5	6.7	158	4 US-09-134-001C-4358	Sequence 4358, Ap
44	70.5	6.7	209	1 US-08-216-593-2	Sequence 2, Appli
45	70.5	6.7	897	4 US-09-134-001C-3600	Sequence 3600, Ap

ALIGNMENTS

RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNGHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475, 886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947, 338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397, 232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match 99.7%, Score 1056, DB 3; Length 2227;
Best Local Similarity 99.5%, Pred. No. 9,8e-113;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEDRPFESHIECRKPYKELRLVGGKQRLKYAOEELSNEVLPPPRKXKGLFSQAK 60
DB 779 VDDPRSEDRPFESHIECRKPYKELRLVGGKQRLKYAOEELSNEVLPPPRKXKGLFSQAK 838
QY 61 ISLFTTEHEIMKPFWRVTAQDTRALRRGFSIAGRGVMTLEMAQVLTGRLINLNEK 120
DB 839 ISLFTTEHEIMKPFWRVTAQDTRALRRGFSIAGRGVMTLEMAQVLTGRLINLNEK 898
QY 121 WTEMKDDIVSLIEKFTSNKTKYSKVPFGMLDLBEIANSKDPFNMSETDLCFLIHLN 180
DB 899 WTEMKDDIVSLIEKFTSNKTKYSKVPFGMLDLBEIANSKDPFNMSETDLCFLIHLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980
RESULT 2
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912

```

; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
; 08-475-886-6

```

```

Query Match          99.7%; Score 1056; DB 3; Length 2227;
Best Local Similarity 99.5%; Pred. No. 9.8e-113;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VDDPRSEEDRRFESHTECKPYKELRLEVKGKORLKYAOEELSNEVLPPPRKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHTECKPYKELRLEVKGKORLKYAOEELSNEVLPPPRKMKGLFSQAK 838
QY 61 ISLFYTEHEHIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEHEHIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKODKIVSLIEKFTSNKYWSKVNPPHGMDLDEEIAANSKDFPMSETDLCFLHMLN 180
DB 899 WTEMKODKIVSLIEKFTSNKYWSKVNPPHGMDLDEEIAANSKDFPMSETDLCFLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

```

```

RESULT 3
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
; 08-397-232-2

```

```

Query Match          99.7%; Score 1056; DB 4; Length 2227;
Best Local Similarity 99.5%; Pred. No. 9.8e-113;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VDDPRSEEDRRFESHTECKPYKELRLEVKGKORLKYAOEELSNEVLPPPRKMKGLFSQAK 60

```

```

DB 779 VDDPRSEEDRRFESHTECKPYKELRLEVKGKORLKYAOEELSNEVLPPPRKMKGLFSQAK 838
QY 61 ISLFYTEHEHIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEHEHIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKODKIVSLIEKFTSNKYWSKVNPPHGMDLDEEIAANSKDFPMSETDLCFLHMLN 180
DB 899 WTEMKODKIVSLIEKFTSNKYWSKVNPPHGMDLDEEIAANSKDFPMSETDLCFLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

```

```

RESULT 4
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
; 08-397-232-4

```

```

Query Match          99.7%; Score 1056; DB 4; Length 2227;
Best Local Similarity 99.5%; Pred. No. 9.8e-113;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VDDPRSEEDRRFESHTECKPYKELRLEVKGKORLKYAOEELSNEVLPPPRKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHTECKPYKELRLEVKGKORLKYAOEELSNEVLPPPRKMKGLFSQAK 838
QY 61 ISLFYTEHEHIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEHEHIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKODKIVSLIEKFTSNKYWSKVNPPHGMDLDEEIAANSKDFPMSETDLCFLHMLN 180
DB 899 WTEMKODKIVSLIEKFTSNKYWSKVNPPHGMDLDEEIAANSKDFPMSETDLCFLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

```

```

RESULT 5
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; APPLICANT: EMERSON, SUZANNE, U.;
; APPLICANT: PURCELL, ROBERT, H.;
; APPLICANT: SIMIAN-HUMAN HAV
; TITLE OF INVENTION: HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:

```



```
ADDRESSSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996

ATTORNEY/AGENT INFORMATION:
NAME: William S. Feiler
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229US1
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match      99.7%; Score 1056; DB 4; Length 2227;
Best Local Similarity 99.5%; Pred. No. 9.8e-113;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRPRESHIECKRPYKELRLVKGKORLKYAOEELSNEVLPPIPRKMKGLFSQAK 60
DB 779 VDDPSEEDKRFESHIIECKRPYKELRLVKGKORLKYAOEELSNEVLPPIPRKMKGLFSQAK 838
QY 61 ISLFYTEHEIMKFSWRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEHEIMKFSWRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDKIVSLIEKFTSNKYWSKVPFHGMGLDEEIAANSKDPFNMSETDLCFLHMLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYWSKVPFHGMGLDEEIAANSKDPFNMSETDLCFLHMLN 958

QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

RESULT 6
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT FILING DATE: US/09/653,499
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
```

```
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match      99.7%; Score 1056; DB 4; Length 2227;
Best Local Similarity 99.5%; Pred. No. 9.8e-113;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRPRESHIECKRPYKELRLVKGKORLKYAOEELSNEVLPPIPRKMKGLFSQAK 60
DB 779 VDDPSEEDKRFESHIIECKRPYKELRLVKGKORLKYAOEELSNEVLPPIPRKMKGLFSQAK 838
QY 61 ISLFYTEHEIMKFSWRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEHEIMKFSWRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDKIVSLIEKFTSNKYWSKVPFHGMGLDEEIAANSKDPFNMSETDLCFLHMLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYWSKVPFHGMGLDEEIAANSKDPFNMSETDLCFLHMLN 958

RESULT 7
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT FILING DATE: US/09/653,499
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match      99.7%; Score 1056; DB 4; Length 2227;
Best Local Similarity 99.5%; Pred. No. 9.8e-113;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRPRESHIECKRPYKELRLVKGKORLKYAOEELSNEVLPPIPRKMKGLFSQAK 60
DB 779 VDDPSEEDKRFESHIIECKRPYKELRLVKGKORLKYAOEELSNEVLPPIPRKMKGLFSQAK 838
QY 61 ISLFYTEHEIMKFSWRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEHEIMKFSWRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDKIVSLIEKFTSNKYWSKVPFHGMGLDEEIAANSKDPFNMSETDLCFLHMLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYWSKVPFHGMGLDEEIAANSKDPFNMSETDLCFLHMLN 958
```

OY 181 PKKINLADRMGLSGVOEIKQ 202
Db 959 PKKINLADRMGLSGVOEIKQ 980

RESULT 8

US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 99.2%; Score 1051; DB 3; Length 2227;
Best Local Similarity 99.0%; Pred. No. 3,7e-112;
Matches 200; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDDPRSEEDRRFESHIECRKPYKELRLEVQKQRLKYAOEELSNEVLPPEPRKMGKGLFSQAK 60
Db 779 VDDPRSEEDRRFESHIECRKPYKELRLEVQKQRLKYAOEELSNEVLPPEPRKMGKGLFSQAK 838
OY 61 ISLFYTEHEHIMKFSWGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
Db 839 ISLFYTEHEHIMKFSWGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
OY 121 WTEMDDKIVSLIEKFTSNKYKSNVPPHGMJLDEIANSKDPFNMSETDLCFLHMLN 180
Db 899 WTEMDDKIVSLIEKFTSNKYKSNVPPHGMJLDEIANSKDPFNMSETDLCFLHMLN 958

181 PKKINLADRMGLSGVOEIKQ 202
959 PKKINLADRMGLSGVOEIKQ 980

RESULT 9

US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 99.2%; Score 1051; DB 4; Length 2227;
Best Local Similarity 99.0%; Pred. No. 3,7e-112;
Matches 200; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDDPRSEEDRRFESHIECRKPYKELRLEVQKQRLKYAOEELSNEVLPPEPRKMGKGLFSQAK 60
Db 779 VDDPRSEEDRRFESHIECRKPYKELRLEVQKQRLKYAOEELSNEVLPPEPRKMGKGLFSQAK 838
OY 61 ISLFYTEHEHIMKFSWGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
Db 839 ISLFYTEHEHIMKFSWGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
OY 121 WTEMDDKIVSLIEKFTSNKYKSNVPPHGMJLDEIANSKDPFNMSETDLCFLHMLN 180
Db 899 WTEMDDKIVSLIEKFTSNKYKSNVPPHGMJLDEIANSKDPFNMSETDLCFLHMLN 958
OY 181 PKKINLADRMGLSGVOEIKQ 202
Db 959 PKKINLADRMGLSGVOEIKQ 980

RESULT 10

5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R., BALTIMORE, DAVID, FEINSTONE,
; STEPHEN M., PURCELL, ROBERT H., RACANIELLO, VINCENT R.,
; BAROUDY, BAHICE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO:2
; LENGTH: 1091
5516630-2

Query Match 37.6%; Score 398; DB 6; Length 1091;
Best Local Similarity 98.7%; Pred. No. 4,5e-37;
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDDPRSEEDRRFESHIECRKPYKELRLEVQKQRLKYAOEELSNEVLPPEPRKMGKGLFSQAK 60
Db 1016 VDDPRSEEDRRFESHIECRKPYKELRLEVQKQRLKYAOEELSNEVLPPEPRKMGKGLFSQAK 1075
OY 61 ISLFYTEHEHIMKFSW 76
Db 1076 ISLFYTEHEHIMKFSW 1091

RESULT 11

US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARCOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1615 L Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,016
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,828
FILING DATE: 03-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 839 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-087-016-2

Query Match 26.2%; Score 277.5; DB 1; Length 839;
Best Local Similarity 85.5%; Pred. No. 2.7e-23;
Matches 53; Conservative 7; Mismatches 1; Indels 1; Gaps 1;
Qy 1 VDDPSEEDRRESHIECKPYKELRLEFGKORLKYAOEELSNEVLPKPKKGLFSQAK 60
Db 779 VDDPTDEDRRESHIE-KKPYKELRLEFGKORFYARELSNEILPPKRLKGLFSQSK 837
Qy 61 IS 62
Db 838 IS 839
BLT 12
US-09-398-193-99
Sequence 99, Application US/09398193
Patent No. 6197581
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: Adenylate cyclase and uses therefor
FILE REFERENCE: P24360-
CURRENT APPLICATION NUMBER: US/09/398,193
CURRENT FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 104
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 99
LENGTH: 1353
TYPE: PRT
ORGANISM: Human
US-09-398-193-99

Query Match 8.0%; Score 84.5; DB 4; Length 1353;
Best Local Similarity 23.5%; Pred. No. 1.2;
Matches 54; Conservative 30; Mismatches 83; Indels 63; Gaps 12;
Qy 16 IECKRPYELRLIE-----VGKORLKYAOEELSNEV-----LPPPKKGLFSQAKISL 63
Db 16 IECKRPYELRLIE-----VGKORLKYAOEELSNEV-----LPPPKKGLFSQAKISL 63

Db 323 LEVEKALKEKRMTHSVMPRIADLIMKOGDESESVKRNHATSSPKNRRKKSSIQAKPIAF 382
Qy 64 --FYTEEHIMKFSWRGVATDTRALRRFGFSLAAGRSVTLTMDAGVLTGRILRLNDEK 121
Db 383 RPFKNQOIEEIVSLFADIVGFTK-----MSAKSHALVGLNLDFGRDRLCE----- 431
Qy 122 TEMKDKIVSLIEKFTSNKYKSNVPH-----GMID-LEBIANSKDPFN 166
Db 432 -ETKCEKISTL-----GDCYVCACCPERRADHAYCCEIMGLGMIKALIEFOCKEKEMYN 485
Qy 167 M-----SETDCEFLH-----WLNPKKILADRM--IGLSGVEIYE 201
Db 486 MKVGVTHTVTLGLIGMRREFDWSN--DVNLNLMIEQLGVAGKVHISE 533

RESULT 13
US-08-245-511-47
Sequence 47, Application US/08245511
Patent No. 5928900
GENERAL INFORMATION:
APPLICANT: Masure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-245-511-47

Query Match 7.8%; Score 82.5; DB 2; Length 643;
Best Local Similarity 26.1%; Pred. No. 0.66;
Matches 42; Conservative 19; Mismatches 53; Indels 47; Gaps 9;

Qy 34 LKYAOEELSNEVLPKPKKGLFSQAKISLFTYEEHIMKFSWRGVTA-DTRALRRFGFS 92
Db 107 LKYADKSDALYPVQESIKGLDAYVK-----GEIKDFQVGIKALDEGTVO---YT 155
Qy 93 LAAGRSVTLTMDAGVLTGRILRLNDEKTEMKD-----KIVSLIEK- 135
Db 156 LNKPSFNSKTTMGVLA---PVNEEFLNSKGDPAKATDPSSLLYNGPYLLKSIYVTS 211

QY 136 ---FTSN-KYWSKV-----PPHGMLEIEIANSKD 163
DB 212 SVEFAKNPYWMDKNDVHIDKVLKSFWDGQ-DTSKPAENFKD 251

RESULT 14

US-08-600-993A-47
Sequence 47, Application US/08600993A
Patent No. 5981229
GENERAL INFORMATION:
APPLICANT: Masure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomenen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,993A
FILING DATE: 1-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-600-993A-47

Query Match 7.8%; Score 82.5; DB 2; Length 643;

Best Local Similarity 26.1%; Pred. No. 0.66; Mismatches 53; Indels 47; Gaps 9;

QY 34 LKAYAEELSNEVLPPPRKMGKLFSSQAKISLPTTEHEIMKFSWGVTA-DTRALRRPFS 92
DB 107 LKAYADKSDALYPVQSIGLDAVVK-----GRTKDSQVIGIKALDEQTVQ---YT 155
QY 93 LAAGRSVWTEMDAGVUTGRLIRLNDKRTMKDD-----KIVSLIEK- 135
DB 156 LNKRESFWSNKTWTGVLA---PVNEEFNLNSKGDGDFAKATDPSSLVNGPYLLKSIYTKS 211
QY 136 ---FTSN-KYWSKV-----PPHGMLEIEIANSKD 163
DB 212 SVEFAKNPYWMDKNDVHIDKVLKSFWDGQ-DTSKPAENFKD 251

RESULT 15

US-09-473-717-2
Sequence 2, Application US/09473717
Patent No. 6372475
GENERAL INFORMATION:
APPLICANT: Storm, Daniel R.
APPLICANT: Hacker, Beth
APPLICANT: Tomlinson, James E.
APPLICANT: COR Therapeutics, Inc.
TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
TITLE OF INVENTION: CYCLASE
FILE REFERENCE: 44481-5029-01-US
CURRENT APPLICATION NUMBER: US/09/473,717
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: PCT/US98/13541
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/098,559
PRIOR FILING DATE: 1997-07-01
PRIOR APPLICATION NUMBER: 08/886,440
PRIOR FILING DATE: 1997-07-01
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1294
TYPE: PRT
ORGANISM: human type IX adenylyl cyclase
US-09-473-717-2

Query Match 7.8%; Score 82.5; DB 4; Length 1294;
Best Local Similarity 23.5%; Pred. No. 1.8; Mismatches 83; Indels 63; Gaps 12;
Matches 54; Conservative 30;

QY 16 IECRPYKELRL-----VGKRLKYAEELSNEV-----LPPPRKMGKLFSSQAKISL 63
DB 323 LEVERALKRMTSHVPRILADDMKQGESESESVKHAATSPNRRKKSSIQAPLAF 382
QY 64 -PYTEHEIMKFSWGVTAADTRALRRFGSLAAGRSVWTEMDAGVUTGRLIRLNDKX 121
DB 383 RPPKQIEVEVSLFADIVGFTK-----MSANSAHALVGLINDLFGFRDRLCE--- 431
QY 122 TEMKDKIVSLKFTSNKYWSKVPPH-----GMLD-LIEIANSKDFPN 166
DB 432 -ETKCEKISTL-----GDCYCVAGCPBRADHAYCCIEMLGMITKALIQFOCKKEMYN 485
QY 167 M-----SETDLCEFLH-----WLPKKNILADRM-LGLSGVQEIKE 201
DB 486 MRVGHVTGIVLGGIIGMRNRFKFDWSN--DVNLANLMEQLGVAGKVIHSE 533

Search completed: April 2, 2003, 11:11:56
Job time : 17.8824 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 11:10:39 ; Search time 22.7315 Seconds
(without alignments)
543.277 Million cell updates/sec

Title: US-10-104-966-12_COPY_779_980
Perfect score: 1059
Sequence: 1 VDDPSEEDRRFRESHIECRK.....KINLADRLGLSGVQIIEKQ 202

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

al number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep: *
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep: *
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep: *
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep: *
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep: *
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep: *
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep: *
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep: *
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep: *
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep: *
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep: *
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep: *
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep: *
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1059	100.0	2227	US-10-104-966-12	Sequence 12, Appl
2	1059	100.0	2227	US-09-929-955-12	Sequence 12, Appl
3	1056	99.7	2227	US-10-135-988-2	Sequence 2, Appl
4	1056	99.7	2227	US-10-135-988-6	Sequence 6, Appl
5	1051	99.2	2227	US-10-135-988-4	Sequence 4, Appl
6	84.5	8.0	1353	US-09-751-1008-99	Sequence 99, Appl
7	83	7.8	501	US-10-066-521-2	Sequence 2, Appl
8	82.5	7.8	1294	US-10-071-223-2	Sequence 2, Appl
9	80.5	7.6	300	US-10-278-173-34	Sequence 34, Appl
10	79.5	7.5	1250	US-09-801-368-164	Sequence 36, Appl
11	79.5	7.5	1353	US-09-751-1008-2	Sequence 2, Appl
12	79.5	7.5	1353	US-10-071-223-3	Sequence 3, Appl
13	79	7.5	1204	US-09-935-291A-9	Sequence 9, Appl
14	77.5	7.3	705	US-09-854-133-186	Sequence 186, App
15	77.5	7.3	705	US-09-738-973-186	Sequence 186, App
16	77	7.3	454	US-09-866-582-8	Sequence 8, Appl
17	76.5	7.2	786	US-09-291-417-23	Sequence 23, Appl
18	76.5	7.2	1001	US-09-291-417-31	Sequence 31, Appl
19	75.5	7.1	297	US-09-815-242-10292	Sequence 10292, A

20	75.5	7.1	297	US-09-815-242-13742	Sequence 13742, A
21	75	7.1	810	US-09-815-242-13636	Sequence 13636, A
22	74	7.0	400	US-09-765-272-190	Sequence 190, App
23	74	7.0	419	US-10-182-263-4	Sequence 4, Appl
24	74	7.0	2000	US-10-010-901-29	Sequence 29, Appl
25	73.5	6.9	748	US-09-291-417-24	Sequence 24, Appl
26	73	6.9	419	US-10-182-263-5	Sequence 5, Appl
27	73	6.9	419	US-10-182-263-6	Sequence 6, Appl
28	71	6.7	419	US-10-182-263-3	Sequence 3, Appl
29	70.5	6.7	476	US-10-147-835-28	Sequence 28, Appl
30	70	6.6	2092	US-10-147-835-12	Sequence 12, Appl
31	69.5	6.6	445	US-10-013-477-13	Sequence 13, Appl
32	69.5	6.6	538	US-09-758-269-8	Sequence 8, Appl
33	69.5	6.6	538	US-09-758-269-18	Sequence 18, Appl
34	69.5	6.6	538	US-09-758-269-33	Sequence 33, Appl
35	69	6.5	478	US-09-866-050A-512	Sequence 512, App
36	69	6.5	860	US-09-824-637-4	Sequence 4, Appl
37	68.5	6.5	364	US-10-198-343-2	Sequence 2, Appl
38	68.5	6.5	364	US-10-197-315-2	Sequence 2, Appl
39	68	6.4	294	US-09-323-998D-4	Sequence 4, Appl
40	68	6.4	316	US-09-815-242-11179	Sequence 11179, A
41	68	6.4	362	US-09-740-288A-34	Sequence 34, Appl
42	68	6.4	363	US-09-740-288A-35	Sequence 35, Appl
43	68	6.4	392	US-09-961-253-3	Sequence 3, Appl
44	68	6.4	628	US-09-881-752A-144	Sequence 144, App
45	68	6.4	870	US-10-078-929-188	Sequence 188, App

ALIGNMENTS

RESULT 1
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.22AUSCI
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Qy	1	VDDPSEEDRRFRESHIECRKPYKELEVGKQRLKYAOEELSNEVLP	100.0%; Score 1059; DB 9; Length 2227; Best local Similarity 100.0%; Pred. No. 3e-103; Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	779	VDDPSEEDRRFRESHIECRKPYKELEVGKQRLKYAOEELSNEVLP	PPRKKKGKGFPSQAK 60
Qy	61	ISLFYTEHEIEMKFSWGVADTRALRRPFGSLAAGRSVWLTMDAGVLTGRLIRLNDEK	120
Db	839	ISLFYTEHEIEMKFSWGVADTRALRRPFGSLAAGRSVWLTMDAGVLTGRLIRLNDEK	898
Qy	121	WTENKDDKIVSLIEKFTSNKYSKKNPFGHMLDEEIAANSQDFPMSTDCFLHMLN	180
Db	899	WTENKDDKIVSLIEKFTSNKYSKKNPFGHMLDEEIAANSQDFPMSTDCFLHMLN	958

QY 181 PKKINLADRMGLSGVOEIQEQ 202
DB 959 PKKINLADRMGLSGVOEIQEQ 980

RESULT 2

US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 1059; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-103;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDRRFESHIECKRPYKELRLEVKGKORLKYAOEELNEVLPPPKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHIECKRPYKELRLEVKGKORLKYAOEELNEVLPPPKMKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRGYADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 120
DB 839 ISLFYTEEHEIMKFSWGRGYADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 898
QY 121 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMDLBEIAANSKDFPMSSETDLCFLIHWLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMDLBEIAANSKDFPMSSETDLCFLIHWLN 958
QY 181 PKKINLADRMGLSGVOEIQEQ 202
DB 959 PKKINLADRMGLSGVOEIQEQ 980

RESULT 3
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 99.7%; Score 1056; DB 9; Length 2227;
Best Local Similarity 99.5%; Pred. No. 6.2e-103;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDRRFESHIECKRPYKELRLEVKGKORLKYAOEELNEVLPPPKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHIECKRPYKELRLEVKGKORLKYAOEELNEVLPPPKMKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRGYADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 120
DB 839 ISLFYTEEHEIMKFSWGRGYADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 898
QY 121 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMDLBEIAANSKDFPMSSETDLCFLIHWLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMDLBEIAANSKDFPMSSETDLCFLIHWLN 958
QY 181 PKKINLADRMGLSGVOEIQEQ 202
DB 959 PKKINLADRMGLSGVOEIQEQ 980

RESULT 4
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 99.7%; Score 1056; DB 9; Length 2227;
Best Local Similarity 99.5%; Pred. No. 6.2e-103;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDRRFESHIECKRPYKELRLEVKGKORLKYAOEELNEVLPPPKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHIECKRPYKELRLEVKGKORLKYAOEELNEVLPPPKMKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRGYADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 120
DB 839 ISLFYTEEHEIMKFSWGRGYADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 898
QY 121 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMDLBEIAANSKDFPMSSETDLCFLIHWLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMDLBEIAANSKDFPMSSETDLCFLIHWLN 958
QY 181 PKKINLADRMGLSGVOEIQEQ 202
DB 959 PKKINLADRMGLSGVOEIQEQ 980

```
RESULT 5
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN M
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match          99.2%; Score 1051; DB 9; Length 2227;
Best Local Similarity 99.0%; Pred. No. 2,1e-102;
Matches 200; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRRSHIECRPKYELRLVKGQRLKYAOEELSNEVLPPrKKKGLFSQAK 60
DB 779 VDDPSEEDRRRSHIECRPKYELRLVKGQRLKYAOEELSNEVLPPrKKKGLFSQAK 838
QY 61 ISLFTTEHEIMKFSMGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFTTEHEIMKFSMGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDKIVSLIEKFTSNKYKSNFPHGMDLEBIAANSKDFPMSETDLCFLHMLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYKSNFPHGMDLEBIAANSKDFPMSETDLCFLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

RESULT 6
US-09-751-100B-99
; Sequence 99, Application US/09751100B
; Patent No. US20020142436A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: Human Adenylate Cyclase and Use Therefor
; FILE REFERENCE: P27948A
; CURRENT APPLICATION NUMBER: US/09/751,100B
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-100B-99

Query Match          8.0%; Score 84.5; DB 10; Length 1353;
Best Local Similarity 23.5%; Pred. No. 4.1;
Matches 54; Conservative 30; Mismatches 83; Indels 63; Gaps 12;

QY 16 IECRPYKELRL-----VGKQRLKYAOEELSNEV-----LPPPRKKKGLFSQAKSL 63
DB 323 IECRPYKELRL-----VGKQRLKYAOEELSNEV-----LPPPRKKKGLFSQAKSL 382
QY 64 --FYTEEHEIMKFSMGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 121
```

```
DB 383 RPFKMOQIEEVSILFADIVGFTK-----MSANKSAHALVGLNDLFGREDRLCE--- 431
QY 122 TEMKDKIVSLIEKFTSNKYKSNFPH-----GMD-LBEIAANSKDFP 166
DB 432 -ETKCEKISTL-----GDCYCVACCPPRADHAYCCIEGMIGMKAIEQFOCKEMYN 485
QY 167 M-----SETDLCFLH-----WLNPKKINLADRM--LGLSGVOEIKE 201
DB 466 MGVGHTRTVLCGILGMRRFKFDVWSN--DVVLAMLMGLGVAAGKVIHSE 533

RESULT 7
US-10-066-521-2
; Sequence 2, Application US/10066521
; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-334001
; CURRENT APPLICATION NUMBER: US/10/066,521
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-066-521-2

Query Match          7.8%; Score 83; DB 9; Length 501;
Best Local Similarity 24.6%; Pred. No. 1.5;
Matches 45; Conservative 34; Mismatches 66; Indels 38; Gaps 11;

QY 6 SEED--RRRSHIECRPKY--EURLVKGQRL-KYAOEELSNEVLPPrKKKGLFS-- 57
DB 51 SKEDVAKLLDKHYPGKQAMWETLNLFLQINRKDLWTKAOEEMRNKLNPKKMKETFOLI 110
QY 58 QAKISLFTTEH--EIMKFSMGRGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLI 114
DB 111 WEKETCLHVPHEFYKETYKNEYKELND-----AYTAARHTVLEGGPDGICKTTL 162
QY 115 R-----LNDEKTEMKDKIV--SLIEKFTSNKYKSNFPHGMDLEBIAANSKDFP 165
DB 163 RKNVMDMABGNLMDKSYNEKLVYRELCSMTTNK-----NFO--ILDWEN--TSLDDP 212
QY 166 NMS 168
DB 213 SLA 215

RESULT 8
US-10-071-223-2
; Sequence 2, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
; APPLICANT: Storm, Daniel R.
; APPLICANT: Hacker, Beth
; APPLICANT: Tomlinson, James E.
; APPLICANT: COR Therapeutics, Inc.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
; TITLE OF INVENTION: CYCLASE
; FILE REFERENCE: 44481-5029-02-US
; CURRENT APPLICATION NUMBER: US/10/071,223
; CURRENT FILING DATE: 2002-02-11
```

PRIOR APPLICATION NUMBER: 09/473,717
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: PCT/US98/13541
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/098,559
PRIOR FILING DATE: 1997-07-01
PRIOR APPLICATION NUMBER: 08/886,440
PRIOR FILING DATE: 1997-07-01
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 1294
TYPE: PRT
ORGANISM: human type IX adenyl cyclase
US-10-071-223-2

Query Match 7.8%; Score 82.5; DB 12; Length 1294;
Best Local Similarity 23.5%; Pred. No. 6.3;
Matches 54; Conservative 30; Mismatches 83; Indels 63; Gaps 12;

16 IECKRKYELRL-----VGKORLKYAOELSNFY-----LPPRRKKGLFSQAKISL 63
Db 323 LEVERALKRMHISVPRRIADLMQGDSESNVYKHAATSSPKNKKKSSIQKAPLAF 382
QY 64 --FYTEHEIMKFSWGVATDTRALRRFGFSIAAGRSVWTLMDAGVLTGRLIRLNDKX 121
Db 383 RPKMQQIEVEVILPADIYGFTR-----MSANSAHALVGLNDLFGFRDLCE--- 431
QY 122 TEMKDKIVSLIEKTSNKYKSKNPPH-----GMLD-LEIIAANSKDPFN 166
Db 432 -ETKCEKISITL-----GDCYVCVACGPEPRADHAYCCIEHGLMIIKAIQFOCEKKEMVN 485
QY 167 M-----SERDLCEHL-----WLPKKINLADRM--IGSGVGEIHE 201
Db 486 MKVGHVHTGVLGILGMRFRKEDVMSN--DVNLANLMEQLGVAGKVHIS 533

RESULT 9

US-10-278-173-34

Sequence 34, Application US/10278173

Publication No. US20030061637A1

GENERAL INFORMATION:

APPLICANT: Jiang, Cai-Zhong

APPLICANT: Broun, Pierre

APPLICANT: Riechmann, Jose-Luis

APPLICANT: Pineda, Omaira

APPLICANT: Zhang, James

APPLICANT: Yu, Guo-Liang

APPLICANT: Pilgrim, Marsha

APPLICANT: Keddle, James

APPLICANT: Heard, Jacqueline

APPLICANT: Reuber, Lynne

APPLICANT: Ratcliffe, Oliver

APPLICANT: Adam, Luc

APPLICANT: Samaha, Raymond

TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION

FILE REFERENCE: MBI-009

CURRENT APPLICATION NUMBER: US/10/278,173

PRIOR FILING DATE: 2002-10-21

PRIOR APPLICATION NUMBER: US/09/533,392

PRIOR FILING DATE: 2000-03-22

PRIOR APPLICATION NUMBER: 60/125,814

PRIOR FILING DATE: 1999-03-23

NUMBER OF SEQ ID NOS: 177

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 34

LENGTH: 300

TYPE: PRT

ORGANISM: Arabidopsis thaliana

FEATURE:

OTHER INFORMATION: G1004

US-10-278-173-34

Query Match 7.6%; Score 80.5; DB 9; Length 300;
Best Local Similarity 23.4%; Pred. No. 1.4;
Matches 44; Conservative 30; Mismatches 87; Indels 27; Gaps 9;

QY 4 PRSEEDRPFESHI--ECKRKYELRLGVKORLKYAOELSNFYLP-----PRMKGLF 56
Db 101 PSNQNQOFEEELSKQIRKPLKISLPKTEWIOFAENTKPEVTKPSEEEKKHYRGV- 159
QY 57 SQAKISLFYTEHEIMKFS---MRGVATD-----RALRRFGFSIAAGRSVWTLMDAGV 108
Db 160 RORPWGKFAAETIRDPNKGSGRWLG-TFDIAIARAAYDEAFLRSGKALINPFLVGVK 218
QY 109 LTGRLIRLNDKEMKDKIVSLIEKTSNKYKSKN-----FPGMDLEIIA-----AN 160
Db 219 WKPRAD--EGEKKRKRDDKVTVEKVLKTEQSVDVNGGTFPFTVNTLTELCDMDLTG 276
QY 161 SKDPPNMS 168
Db 277 FLNFPILS 284

RESULT 10

US-09-801-368-364

Sequence 364, Application US/09801368

Patent No. US20020128250A1

GENERAL INFORMATION:

APPLICANT: Busby, Robert

APPLICANT: Call, Brian

APPLICANT: Hecht, Peter

APPLICANT: Holtzman, Doug

APPLICANT: Madden, Kevin

APPLICANT: Maxon, Mary

APPLICANT: Milne, Todd

APPLICANT: No. US20020128250A1man, Thea

APPLICANT: Royer, John

APPLICANT: Salama, Sofie

APPLICANT: Sherman, Amir

APPLICANT: Silva, Jeff

APPLICANT: Summers, Eric

TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

FILE REFERENCE: 109272.147

CURRENT APPLICATION NUMBER: US/09/801,368

CURRENT FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: US 09/487,558

PRIOR FILING DATE: 2000-01-19

PRIOR APPLICATION NUMBER: US 60/160,587

PRIOR FILING DATE: 1999-10-20

NUMBER OF SEQ ID NOS: 440

SOFTWARE: Patentin version 3.0

SEQ ID NO 364

LENGTH: 1250

TYPE: PRT

ORGANISM: Saccharomyces cerevisiae

US-09-801-368-364

Query Match 7.5%; Score 79.5; DB 10; Length 1250;
Best Local Similarity 21.0%; Pred. No. 12;
Matches 52; Conservative 28; Mismatches 65; Indels 103; Gaps 12;

QY 41 LNEVLP--PRMKGLFSQAKI--SLFY-----TEHEIMKFS--W----- 76
Db 659 LNEVLDQKNQPKERKSFQPLPLTAESLEYRRNFTDYNENIFAISLGWSEFALHVN 718
QY 77 -----RGVATDTRALRRFG-----FSLAAGR 97
Db 719 NGNGTLELGHVAVDTVSHIEBSSVDRARKRSAAVFMQKLVNLLPQSFNDELSLAPGK 778
QY 98 S-----VWTLK-----DAGVLTGRLIRLN--DEKTEMKDKIVSLIEKTS 138
Db 779 ESATLSVVTTLDSSTIRKSTWGEESTISPSNLSLEQDLKSTGSPSYSTVGEIAR 838
QY 139 NKYKSKNPPHGMLEIIAANSKDPFNSETDLCFLHWNLPKTI---NLADRMGLIS 194


```
Db 839 SFYARINDPEKTL-----LPTLS-----LLESDEKKYVDNLIDRTIGV 881
Qy 195 GVOEIKQ 202
Db 882 VINEIKRK 889

RESULT 11
US-09-751-100B-2
; Sequence 2, Application US/09751100B
; Patent No. US20020142436A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor
; FILE REFERENCE: P27948A
; CURRENT APPLICATION NUMBER: US/09/751,100B
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: Mouse
US-09-751-100B-2

Query Match 7.5%; Score 79.5; DB 10; Length 1353;
Best Local Similarity 23.0%; Pred. No. 14;
Matches 53; Conservative 31; Mismatches 83; Indels 63; Gaps 12;

Qy 16 IECKRPYKELRL-----VGKQRLKYAOEELSNEY-----LPPPRKKKGFPQAKISL 63
Db 323 LEVEALKERMHVSMPRIADLKKQGESESVKHAATSSPKNRKKSSIOKAPIAF 382
Qy 64 --FYTEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDKX 121
Db 393 RPFKMQOIEEVSILPADIYGFTR-----MSANKSAHALVGLNDLFGFRDLCE--- 431
Qy 122 TEMKDKIVSLIEKFTSNKYMSKNVPH-----GMD-LBEIANSKDPFN 166
Db 432 -QTKCEKISTL-----GDCYCVAGCPPEPRADHAYCCIEMGIMTKALIQFOEKKEMVN 485
Qy 167 M-----SETDLCFLH-----WLNPKKINLADRM--LGLSGVOEIKE 201
Db 486 MRVGVHTGTGCGILGMRRFRKFDVWSN--DVNLAMLMEDQLGVAGKVHISE 533

RESULT 12
US-10-071-223-3
; Sequence 3, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
; APPLICANT: Storm, Daniel R.
; APPLICANT: Hacker, Beth
; APPLICANT: Tomlinson, James E.
; APPLICANT: COR Therapeutics, Inc.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
; FILE REFERENCE: 44481-5029-02-US
; CURRENT APPLICATION NUMBER: US/10/071,223
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 09/473,717
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PCT/US98/13541
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/098,559
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 08/886,440
; PRIOR FILING DATE: 1997-07-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1353
```

```
; TYPE: PRT
; ORGANISM: murine type IX adenylyl cyclase
US-10-071-223-3

Query Match 7.5%; Score 79.5; DB 12; Length 1353;
Best Local Similarity 18.4%; Pred. No. 13;
Matches 39; Conservative 36; Mismatches 75; Indels 62; Gaps 8;

Qy 16 IECKRPYKELRL-----VGKQRLKYAOEELSNEY-----LPPPRKKKGFPQAKISL 63
Db 323 LEVEALKERMHVSMPRIADLKKQGESESVKHAATSSPKNRKKSSIOKAPIAF 382
Qy 64 --FYTEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDKX 121
Db 393 RPFKMQOIEEVSILPADIYGFTR-----MSANKSAHALVGLNDLFGFRDLCE--- 431
Qy 122 TEMKDKIVSLIEKFTSNKYMSKNVPH-----GMD-LBEIANSKDPFN 166
Db 432 -QTKCEKISTL-----GDCYCVAGCPPEPRADHAYCCIEMGIMTKALIQFOEKKEMVN 485
Qy 167 M-----SETDLCFLH-----WLNPKKINLADRM--LGLSGVOEIKE 201
Db 486 MRVGVHTGTGCGILGMRRFRKFDVWSN--DVNLAMLMEDQLGVAGKVHISE 533

RESULT 13
US-09-935-291A-9
; Sequence 9, Application US/09935291A
; Patent No. US20020107373A1
; GENERAL INFORMATION:
; APPLICANT: Chun, Miyoung
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 49937, 49931, AND 49933, NOVEL HUMAN TRANSPORTER
; FILE REFERENCE: MNI-184
; CURRENT APPLICATION NUMBER: US/09/935,291A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,504
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: 60/250,932
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1204
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-291A-9

Query Match 7.5%; Score 79; DB 10; Length 1204;
Best Local Similarity 18.4%; Pred. No. 13;
Matches 39; Conservative 36; Mismatches 75; Indels 62; Gaps 8;

Qy 1 VDDPSEEDRRPRESHIECKRPYKELRLVEGQRLKYAOEELSNEYLPPPRKK----- 53
Db 464 IEGTKDPRNRKYKLPLECTL-----ILTSVPPPLPIELSLAVNT 503
Qy 54 GIFSQAQKISLFTTE-----EHEIMKFSWGR-VTADTRALRRFGSLAAGRSVW---- 100
Db 504 SLIALAKLYMCTEPRRIIPFAGKVEVCCFDKGTGLTSLLVVRGVA--GLRDKETVTPVSS 562
Qy 101 -----TLMDAGVLTGRLIR--LNDEKTEMKDKIVSLIEKFTSNKYMS 143
Db 563 IPVETHRALASCHSLMQDDGTGLVDGDPLEKAMLTAVDWTLTVDENVFPRSSITGGLKIHQ 622
Qy 144 KVNPFHGMIDLEEIANSKDPFNMSETDLCFL 175
Db 623 RPFASALKRMSVLSAYEK-----LGSIDLCTYI 650

RESULT 14
US-09-854-133-186
; Sequence 186, Application US/09854133
```

Publication No. US20020183499A1
GENERAL INFORMATION:
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamath, Raedoh
APPLICANT: Henderson, Robert A.
APPLICANT: Benson, Darin R.
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITL OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.475C10
CURRENT APPLICATION NUMBER: US/09/854,133
CURRENT FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 735
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 186
LENGTH: 705
TYPE: PRT
ORGANISM: Homo sapien
US-09-854-133-186

Query Match 7.3%; Score 77.5; DB 9; Length 705;
Best Local Similarity 23.9%; Pred. No. 9.2;
Matches 53; Conservative 18; Mismatches 84; Indels 67; Gaps 9;

QY 2 DDPRSEEDRRFESHIECKRPYKELREVGKORLKYAOEELS-NEVLPPPRKMKGLFSQAK 60
DB 151 DDEPEDEDEEMGHAEIYAEIYMPKIKIGLRHPDAVETSSLSVTTPDVWYKTSISEET 210
QY 61 IS-----LFYTEHEIMKFSWRGVTAADTRALRRFGFSIAAGRSVWTLMDAGV 108
DB 211 IDNGWLSALQLEAIYAAQOHE--TFLPNG-----DRAGFLIGDG-----AGV 251
QY 109 LTGRLLR-----LNDEKWTMKD-----DKIYSLIEKFTSNKYW 142
DB 252 GKGRITAGIYENYLLSRKRALMFSVSNLDKXDAERDLRDIGAKNILLVHSLNKFKYKIS 311
QY 143 SKVNFPHGMLDLEIYAANSKDFPNMSE-----TDLCFLLHW 178
DB 312 SKHN---GSVKKGVIFATYSSLIGESGSGKKYKTRLKQLLHW 350

RESULT 15
US-09-738-973-186
Sequence 186, Application US/09738973
Patent No. US20020110563A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Henderson, Robert A.
APPLICANT: Lodes, Michael J.
APPLICANT: Fling, Steven P.
APPLICANT: Mohamath, Raedoh
APPLICANT: Algate, Paul A.
APPLICANT: Secrist, Heather
APPLICANT: Indirias, Carol Yoseph
APPLICANT: Benson, Darin R.
APPLICANT: Elliot, Mark
APPLICANT: Mannion, Jane
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITL OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.475C9
CURRENT APPLICATION NUMBER: US/09/738,973
CURRENT FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 587
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 186
LENGTH: 705
TYPE: PRT
ORGANISM: Homo sapien
US-09-738-973-186

Query Match 7.3%; Score 77.5; DB 10; Length 705;
Best Local Similarity 23.9%; Pred. No. 9.2;

Matches 53; Conservative 18; Mismatches 84; Indels 67; Gaps 9;
QY 2 DDPRSEEDRRFESHIECKRPYKELREVGKORLKYAOEELS-NEVLPPPRKMKGLFSQAK 60
DB 151 DDEPEDEDEEMGHAEIYAEIYMPKIKIGLRHPDAVETSSLSVTTPDVWYKTSISEET 210
QY 61 IS-----LFYTEHEIMKFSWRGVTAADTRALRRFGFSIAAGRSVWTLMDAGV 108
DB 211 IDNGWLSALQLEAIYAAQOHE--TFLPNG-----DRAGFLIGDG-----AGV 251
QY 109 LTGRLLR-----LNDEKWTMKD-----DKIYSLIEKFTSNKYW 142
DB 252 GKGRITAGIYENYLLSRKRALMFSVSNLDKXDAERDLRDIGAKNILLVHSLNKFKYKIS 311
QY 143 SKVNFPHGMLDLEIYAANSKDFPNMSE-----TDLCFLLHW 178
DB 312 SKHN---GSVKKGVIFATYSSLIGESGSGKKYKTRLKQLLHW 350

Search completed: April 2, 2003, 11:22:16
Job time : 27.7315 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 11:05:19 ; Search time 27.3811 Seconds
(without alignments)
709.218 Million cell updates/sec

Title: US-10-104-966-12_COPY_779_980

Perfect score: 1059
Sequence: 1 VDDPSEEDRRFRESHIECRK.....KINLADRMGLSGVQEIKEQ 202

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Final number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1059	100.0	2227	1 GNNYHR	genome polypotein
2	1056	99.7	2227	1 GNNYHM	genome polypotein
3	1051	99.2	2227	1 GNNYMK	genome polypotein
4	1050	99.2	2227	1 GNNYHB	genome polypotein
5	971	91.7	2230	1 GNNYSA	genome polypotein
6	399	37.7	1358	1 A03905	genome polypotein
7	386	36.4	852	1 GNNYHA	genome polypotein
8	305	28.8	341	2 S04137	genome polypotein
9	277.5	26.2	839	1 GNNYS2	genome polypotein
10	213	20.1	56	2 P00427	genome polypotein
11	212	20.0	56	2 P00429	genome polypotein
12	208	19.6	56	2 P00430	genome polypotein
13	205	19.4	56	2 P00432	genome polypotein
14	204	19.3	56	2 P00428	genome polypotein
15	203	19.2	56	2 P00431	genome polypotein
16	200	18.9	56	2 P00434	genome polypotein
17	189	17.8	55	2 P00433	genome polypotein
18	180	17.0	56	2 P00436	genome polypotein
19	179	16.9	56	2 P00437	genome polypotein
20	178.5	16.9	55	2 P00435	genome polypotein
21	168	15.9	319	2 JH0135	genome polypotein
22	93	8.8	330	2 A83267	hypothetical prote
23	90.5	8.5	309	2 H71089	hypothetical prote
24	89	8.4	1289	1 RMXR3	mRNA guanylyltrans
25	88	8.3	198	2 T09743	hypothetical protei
26	87.5	8.3	235	2 G90224	peptidase M1 fam11
27	86.5	8.2	588	2 E87521	hypothetical prote
28	85	8.0	381	2 T18876	hypothetical prote
29	85	8.0	723	2 T49985	oligopeptidase A-1

30	84	7.9	570	2 A45249	alpha-glucosidase
31	83.5	7.9	610	2 T06690	galactonolactone d
32	83	7.8	355	2 A70185	conserved hypothet
33	82.5	7.8	643	2 S54395	permease-like prot
34	82	7.7	278	2 B82388	probable maltose o
35	82	7.7	387	2 G87537	acyl-CoA dehydroge
36	81	7.6	500	2 C75455	carboxypeptidase-r
37	81	7.6	529	2 G90286	hypothetical prote
38	81	7.6	582	2 S19424	hypothetical prote
39	80.5	7.6	300	2 T52020	ethylene responsiv
40	80.5	7.6	1034	2 S50917	aminomethyltransfe
41	80.5	7.6	1225	2 A49464	chromosome segrega
42	80	7.6	743	2 C6168	hypothetical prote
43	80	7.6	1116	2 H97970	type I site-specif
44	79.5	7.5	387	2 H75009	hypothetical prote
45	79.5	7.5	1006	2 S20126	exoribonuclease RA

ALIGNMENTS

RESULT 1

GNNYHR
genome polypotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Naajarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nest
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.
A:Reference number: A03903; MUID:85190549; PMID:2966127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NA>
A:Cross-references: GB:K02990; NID:G32956; PIDN:AAA45472.1; PID:G329597
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleic acidyltrans
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 1059; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 9.6e-84; Indels 0; Gaps 0;

Matches 202; Conservative 0; Mismatches 0;

QY	1	VDDPSEEDRRFRESHIECRKPYKELRLEVGKORLKYAOEISNEVLPPEPRKMKGLFSQAK	60
DB	779	VDDPSEEDRRFRESHIECRKPYKELRLEVGKORLKYAOEISNEVLPPEPRKMKGLFSQAK	838
QY	61	ISLFTTEHEIHKFSWGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDKX	120
DB	839	ISLFTTEHEIHKFSWGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDKX	898
QY	121	WTMMDKDKIVSLIEKFTSNKYSKNPFHGMIDLEISANSKDFPMWSETDLCFLILHWIN	180
DB	899	WTMMDKDKIVSLIEKFTSNKYSKNPFHGMIDLEISANSKDFPMWSETDLCFLILHWIN	958
QY	181	PKKINLADRMGLSGVQEIKEQ 202	
DB	959	PKKINLADRMGLSGVQEIKEQ 980	

RESULT 2

GNMYHM
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1998 #sequence_revision 31-Mar-1998 #text_change 16-Jul-1999
C:Accession: A25981
R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J:Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with da
A:Reference number: A25981; MUID:87061253; PMID:3023706
A:Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M14707; NID:g329582; PTDN:AAA45465.1; PID:g329583
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-23/Product: coat protein 1A #status predicted <VP1>
F:1-245/Product: coat protein 1B #status predicted <VP2>
F:1-46-491/Product: coat protein 1C #status predicted <VP3>
F:1492-791/Product: coat protein 1D #status predicted <VP1>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 99.7%; Score 1056; DB 1; Length 2227;
Best Local Similarity 99.5%; Pred. No. 1.8e-83;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDRRFESHIECKRPYKELRLEVGNKORLKYAOEELSNEVLPPEPRKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHIECKRPYKELRLEVGNKORLKYAOEELSNEVLPPEPRKMKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRVADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEEHEIMKFSWGRVADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDDKIVSLIEKFTSNKYKSKNPFPHGMLDLEIANSKDPNNMSETDLCFLHMLN 180
DB 899 WTEMKDDKIVSLIEKFTSNKYKSKNPFPHGMLDLEIANSKDPNNMSETDLCFLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

RESULT 3
GNMYHM
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
A:Reference number: A94149; MUID:8715701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PTDN:AAA45471.1; PID:g329595
A:Note: Submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-245/Product: coat protein 1A #status predicted <P1A>

F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: core protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 99.2%; Score 1051; DB 1; Length 2227;
Best Local Similarity 99.0%; Pred. No. 4.8e-83;
Matches 200; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDRRFESHIECKRPYKELRLEVGNKORLKYAOEELSNEVLPPEPRKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHIECKRPYKELRLEVGNKORLKYAOEELSNEVLPPEPRKMKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRVADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEEHEIMKFSWGRVADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDDKIVSLIEKFTSNKYKSKNPFPHGMLDLEIANSKDPNNMSETDLCFLHMLN 180
DB 899 WTEMKDDKIVSLIEKFTSNKYKSKNPFPHGMLDLEIANSKDPNNMSETDLCFLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

RESULT 4
GNMYHM
genome polyprotein - human hepatitis A virus (strain MB8)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
Vpg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klein, R.; Wimmer, E.; Deinhardt,
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PNU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydr
F:1-24/Product: coat protein 1A #status predicted <VP1>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein Vpg #status predicted <VP3>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 99.2%; Score 1050; DB 1; Length 2227;
Best Local Similarity 99.0%; Pred. No. 5.9e-83;
Matches 200; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDDPRSEEDRRFESHIECKRPYKELRLEVGNKORLKYAOEELSNEVLPPEPRKMKGLFSQAK 60
DB 779 VDDPRSEEDRRFESHIECKRPYKELRLEVGNKORLKYAOEELSNEVLPPEPRKMKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRVADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEEHEIMKFSWGRVADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898

Qy 121 WTEKDDKIVSLIEKFTSNKTKSVNPPHGMIDLEELIANSKDPNNSETDLCFILHMLN 180
Db 899 WTEKDDKIVSLIEKFTSNKTKSVNPPHGMIDLEELIANSKDPNNSETDLCFILHMLN 958
Qy 181 PKKINLADRMGLSGVGEIKEQ 202
Db 959 PKKINLADRMGLSGVGEIKEQ 980

RESULT 5

GENYSA
genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S0485; S03965
R:Tsarev, S.A.
Submitted to JIPID, April 1991
A:Reference number: A30470
Molecule type: genomic RNA

A:Residues: 1-2230 <TSA>
A:Cross-References: GB:D00924; NID:G222597; PIDN:BA00766.1; PID:G222598
R:Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure at
A:Reference number: J01080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885

A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL2>
A:Cross-References: EMBL:X15461; NID:G61971; PIDN:CA033490.1; PID:G9330268
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he
A:Reference number: S03965; MUID:85232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-References: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polypeptide
F:1-27/Product: coat protein 1A #status predicted <C1A>
F:28-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: core protein 3A #status predicted <C3A>
F:1499-1521/Product: core protein 3B #status predicted <C3B>
F:1522-1741/Product: core protein 3C #status predicted <C3C>
F:1742-2230/Product: protein 3D #status predicted <C3D>

Query Match 91.7%; Score 971; DB 1; Length 2230;
Best Local Similarity 89.6%; Pred. No. 4.5e-76;
Matches 181; Conservative 15; Mismatches 6; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRRPESHIECRKPYKELRLVGGKORLKYAOEELSNEVLPPPRKKGGLFSQAK 60
Db 783 VDDPSEADKPFESHIECKPKYKELRLVGGKORLKYAOEELSNEVLPPPRKKGGLFSQAK 842
Qy 61 ISLFTTEHEIMKFSWRCVLTADTRALRRFGFSIAGRSVWTLMDAGVLTGLIRLNDK 120
Db 843 ISLFTTEHEIMKFSWRCVLTADTRALRRFGFSIAGRSVWTLMDAGVLTGLIRLNDK 902
Qy 121 WTEKDDKIVSLIEKFTSNKTKSVNPPHGMIDLEELIANSKDPNNSETDLCFILHMLN 180
Db 959 WTEKDDKIVSLIEKFTSNKTKSVNPPHGMIDLEELIANSKDPNNSETDLCFILHMLN 958

Db 903 WTEKDDKIVSLIEKFTSNKTKSVNPPHGMIDLEELIANSKDPNNSETDLCFILHMLN 962
Qy 181 PKKINLADRMGLSGVGEIKEQ 202
Db 963 PKKINLADRMGLSGVGEIKEQ 984

RESULT 6

A03905
genome polyprotein (version 2) - human hepatitis A virus (fragments)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein
C:Species: human hepatitis A virus
C:Date: 31-Mar-1991 #sequence revision 31-Mar-1991 #text_change 15-Nov-1996
C:Accession: A03905
R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinston
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985
A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A:Reference number: A03905; MUID:85166289; PMID:2984684
A:Accession: A03905
A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polypeptide
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>
F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 37.7%; Score 399; DB 2; Length 1358;
Best Local Similarity 98.7%; Pred. No. 1.8e-26;
Matches 75; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRRPESHIECRKPYKELRLVGGKORLKYAOEELSNEVLPPPRKKGGLFSQAK 60
Db 779 VDDPSEADKPFESHIECKPKYKELRLVGGKORLKYAOEELSNEVLPPPRKKGGLFSQAK 838
Qy 61 ISLFTTEHEIMKFSW 76
Db 839 ISLFTTEHEIMKFSW 854

RESULT 7

GENYHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W
J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648; PMID:2985793
A:Accession: A03904
A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>
A:Cross-References: EMBL:M10033; NID:G329592; PIDN:AAA5470.1; PID:G329593
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polypeptide
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 36.4%; Score 386; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.4e-25;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRRPESHIECRKPYKELRLVGGKORLKYAOEELSNEVLPPPRKKGGLFSQAK 60
Db 779 VDDPSEADKPFESHIECKPKYKELRLVGGKORLKYAOEELSNEVLPPPRKKGGLFSQAK 838

OY 61 ISLFYTEHEIMKF 74
|||||
Db 839 ISLFYTEHEIMKF 852

RESULT 8

genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Antonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
A:Reference number: S04137; PMID:89263805; PMID:2542903
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:G62301; PIDN:CAA32794.1; PID:G4377576
A:Genetics:
A:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:2-340/Product: coat protein ID (VP1) #status predicted <MAT>

Query Match 28.8%; Score 305; DB 2; Length 341;
Best Local Similarity 98.3%; Pred. No. 5e-19;
Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEELNPPPKMKGLFSQA 59
|||||
Db 283 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEELNPPPKMKGLFSQS 341

RESULT 9

genome polyprotein - simian hepatitis A virus (strain CY-145) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat protein
C:Species: simian hepatitis A virus
A:Note: host Macaca fascicularis (Cynomolgus macaque)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
C:Accession: JQ1180
R:Nainan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.
J. Gen. Virol. 72, 1685-1689, 1991
A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus mac
A:Reference number: JQ1180; PMID:91311421; PMID:1649902
A:Accession: JQ1180
A:Molecule type: genomic RNA
A:Residues: 1-839 <NA1>

C:Cross-references: GB:M59286; NID:G329599; PIDN:AAA45473.1; PID:G555083
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-23/Product: coat protein 1A #status predicted <VP0>
F:24-245/Product: coat protein 1B #status predicted <VP3>
F:246-491/Product: coat protein 1C #status predicted <VP1>
F:492-839/Product: core protein 2A (fragment) #status predicted <P2P>
F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.2%; Score 277.5; DB 1; Length 839;
Best Local Similarity 85.5%; Pred. No. 3.9e-16;
Matches 53; Conservative 7; Mismatches 1; Indels 1; Gaps 1;

OY 1 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEELNPPPKMKGLFSQA 60
|||||
Db 779 VDDPRTDEDRFRESHIE-KRPYKELRLEVKGKORLKYARELSNEIILPPPKMLGLFSQSK 837

OY 61 IS 62
||
Db 838 IS 839

RESULT 10

PQ0427

genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0427
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel

J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; PMID:92300330; PMID:1318940
A:Accession: PQ0427
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>

C:Comment: This protein is from the VP1/2A junction region.
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 20.1%; Score 213; DB 2; Length 56;
Best Local Similarity 97.6%; Pred. No. 5.5e-12;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEEL 41
|||||
Db 16 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEEL 56

RESULT 11

genome polyprotein - human hepatitis A virus (strain PRCL6) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0429
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; PMID:92300330; PMID:1318940
A:Accession: PQ0429
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>

C:Note: this protein is from the VP1/2A junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 20.0%; Score 212; DB 2; Length 56;
Best Local Similarity 97.6%; Pred. No. 6.7e-12;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEEL 41
|||||
Db 16 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEEL 56

RESULT 12

genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0430
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; PMID:92300330; PMID:1318940
A:Accession: PQ0430
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>

A:Note: this protein is from the VP1/2A junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 19.6%; Score 208; DB 2; Length 56;
Best Local Similarity 97.6%; Pred. No. 1.5e-11;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VDDPSEEDRRFRESHIECKRPYKELRLEVKGKORLKYAOEEL 41
|||||

Db 16 VDDPRSEEDRRPESHIECKRPYKELRLEVGVKQRLKYAOEEL 56

RESULT 13

PO0432 genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)

C/Species: human hepatitis A virus

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: P00432

R/Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel

J. Gen. Virol. 73, 1365-1377, 1992

A/Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr

A/Reference number: P00427; PMID:92300330; PMID:1318940

A/Accession: P00432

A/Molecule type: mRNA

A/Residues: 1-56 <ROB>

A/Note: this protein is from the VP1/2A junction region

C/Superfamily: hepatitis A virus genome polyprotein

C/Keywords: coat protein; core protein; polyprotein

Query Match Best Local Similarity 19.4%; Score 205; DB 2; Length 56;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDRRPESHIECKRPYKELRLEVGVKQRLKYAOEEL 41
Db 16 VDDPRSEEDRRPESHIECKRPYKELRLEVGVKQRLKYAOEEL 56

RESULT 14

PO0428 genome polyprotein - human hepatitis A virus (strain TKW002) (fragment)

C/Species: human hepatitis A virus

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: P00428

R/Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel

J. Gen. Virol. 73, 1365-1377, 1992

A/Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr

A/Reference number: P00427; PMID:92300330; PMID:1318940

A/Accession: P00428

A/Molecule type: mRNA

A/Residues: 1-56 <ROB>

A/Note: this protein is from the VP1/2A junction region

C/Superfamily: hepatitis A virus genome polyprotein

C/Keywords: coat protein; core protein; polyprotein

Query Match Best Local Similarity 19.3%; Score 204; DB 2; Length 56;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDRRPESHIECKRPYKELRLEVGVKQRLKYAOEEL 41
Db 16 VDDPRSEEDRRPESHIECKRPYKELRLEVGVKQRLKYAOEEL 56

RESULT 15

PO0431 genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)

C/Species: human hepatitis A virus

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: P00431

R/Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel

J. Gen. Virol. 73, 1365-1377, 1992

A/Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr

A/Reference number: P00427; PMID:92300330; PMID:1318940

A/Accession: P00431

A/Molecule type: mRNA

A/Residues: 1-56 <ROB>

A/Note: this protein is from the VP1/2A junction region

C/Superfamily: hepatitis A virus genome polyprotein

C/Keywords: coat protein; core protein; polyprotein

Query Match 19.2%; Score 203; DB 2; Length 56;

Best Local Similarity 95.1%; Pred. No. 4.1e-11;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDRRPESHIECKRPYKELRLEVGVKQRLKYAOEEL 41
Db 16 VDDPRSEEDRRPESHIECKRPYKELRLEVGVKQRLKYAOEEL 56

Search completed: April 2, 2003, 11:11:23
Job time : 30.381 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.ps_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 11:03:29 ; Search time 14.4655 Seconds
(without alignments)
579,187 Million cell updates/sec

Title: US-10-104-966-12_COPY_779_980
Perfect score: 1059
Sequence: 1 VDDPRSEDRRFRSHIECRK.....KINLADRMGLSGVQEIKEQ 202

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1059	100.0	2227 1	POLG_HPAVL
2	1056	99.7	2227 1	POLG_HPAVL
3	1050	99.2	2227 1	POLG_HPAVL
4	1048	99.0	2226 1	POLG_HPAVL
5	1048	99.0	2226 1	POLG_HPAVL
6	1044	98.6	2226 1	POLG_HPAVL
7	971	91.7	2230 1	POLG_HPAVL
8	386	36.4	852 1	POLG_HPAVL
9	305	28.8	341 1	POLG_HPAVL
10	277.5	26.2	839 1	POLG_HPAVL
11	251	23.7	808 1	POLG_HPAVL
12	89	8.4	1289 1	MCE_RBOVD
13	85	8.0	297 1	RM35_MOUSE
14	85	8.0	381 1	RIR2_CAEBL
15	84.5	8.0	1353 1	CV99_HUMAN
16	84	7.9	569 1	MALT_CANAL
17	83	7.8	355 1	TRMU_HOBU
18	81.5	7.7	976 1	SCPI_HUMAN
19	81	7.6	581 1	DPO4_YEAST
20	80.5	7.6	300 1	ERF5_ARATH
21	80.5	7.6	1034 1	GSP_YEAST
22	80.5	7.6	1225 1	SMC1_YEAST
23	80	7.6	1200 1	ATY2_MOUSE
24	79.5	7.5	1006 1	RAT1_YEAST
25	79.5	7.5	1250 1	SDD1_YEAST
26	79.5	7.5	1353 1	CTA9_MOUSE
27	79.5	7.5	1374 1	REPOB_RICPR
28	79	7.5	739 1	RELA_STREO
29	79	7.5	1404 1	ATY2_HUMAN
30	77.5	7.3	1440 1	SYEP_HUMAN
31	77	7.3	242 1	TPM1_PODCA
32	76.5	7.2	364 1	MKI1_MOUSE
33	76.5	7.2	370 1	MPK3_ARATH

34	76.5	7.2	381 1	STW_SULTO	Q976m1	bulfolobus
35	76.5	7.2	1099 1	PLCI_CANAL	Q13433	candida alb
36	75.5	7.1	297 1	ICIA_ECOLI	P24194	escherichia
37	75.5	7.1	297 1	ICIA_SALTI	P58508	salmonella
38	75.5	7.1	297 1	ICIA_SALTY	P58509	salmonella
39	75.5	7.1	971 1	USG1_MOUSE	O08810	mus musculus
40	75.5	7.1	972 1	USG1_HUMAN	Q15029	homo sapien
41	75	7.1	628 1	DY11_MOUSE	O88485	mus musculus
42	75	7.1	643 1	DY11_RAT	O63100	rattus norv
43	75	7.1	645 1	DY11_HUMAN	O14576	homo sapien
44	75	7.1	729 1	KAR3_YEAST	P17119	saccharomyc
45	75	7.1	1043 1	ER3A_YEAST	P16521	saccharomyc

ALIGNMENTS

RESULT 1
POLG_HPAVL STANDARD; PRT; 2227 AA.
ID POLG_HPAVL
AC P06441;
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein (Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (RC 2.7.7.48)).
OS Hepatitis A virus (strain LA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NC NCB1 Taxid=12099;
RX [1]
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Merryweather J., van Nest G., Dima D.;
RT "Primary structure and gene organization of human hepatitis A virus";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; K02990; AAA5472.1; -.
CC PIR; A03903; GNNYHR.
CC MEROPS; C03.005; -.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR001205; RNA_pol_P3D.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Core protein; Transferrase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
PROBABLE PROTEIN VP4 (P1A).
COAT PROTEIN VP2 (P1B).
COAT PROTEIN VP3 (P1C).
COAT PROTEIN VP1 (P1D).
CORE PROTEIN P2A.
CORE PROTEIN P2B.
CORE PROTEIN P2C.
PROBABLE PROTEIN P3A.
PROBABLE PROTEIN P3B.
PROBABLE PROTEIN P3C.

```

FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 1059; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-87;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRREESHIECKRPYKELELVGKORLKYAQBELSENVLPPEPKKKGLFSQAK 60
Db 779 VDDPSEEDRREESHIECKRPYKELELVGKORLKYAQBELSENVLPPEPKKKGLFSQAK 838

Qy 61 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRINDEX 120
Db 839 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRINDEX 898

Qy 121 WTEMODKIVSLIEKTSKWKYSKNVPHGMDLBEIAANSKDFPMSTDLCPFLHWIN 180
Db 899 WTEMODKIVSLIEKTSKWKYSKNVPHGMDLBEIAANSKDFPMSTDLCPFLHWIN 958

Db 181 PKKINLADRLGLSGVQEIKEQ 202
959 PKKINLADRLGLSGVQEIKEQ 980

RESULT 2
POLG_HPAVH STANDARD; PRT; 2227 AA.
ID POLG_HPAVH 081082;
AC P08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08; Created)
DT 01-AUG-1988 (Rel. 08; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Hepatitis A virus; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepacovirus.
OX NCBI_TaxID=12098;
RN [1]
RP STRAIN=Wild type;
RC MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A., Barclay B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses.";
RT J. Virol. 61:50-59 (1987).
[2]
RC SEQUENCE FROM N.A. STRAIN=Attenuated;
RA MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M., Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with wild-type virus.";
RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501 (1987).
[3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RA MEDLINE=85166289; PubMed=2984684;
RA Barclay B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr., Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA polymerase.";
RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147 (1985).
RC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.

```

```

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M14114; AAA45475.1; -.
CC EMBL; M14707; AAA45465.1; -.
CC EMBL; M16632; AAA45471.1; -.
CC PIR; A25981; GNNYTM.
CC PIR; A25914; GNNYMK.
CC MEROPS; C03.005; -.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR001205; RNA_pol_P3D.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC RNA-directed RNA polymerase; Core protein; Core protein; Transferase; Hydrolase; Thiol protease.
CC CHAIN 1 23
CC CHAIN 24 245
CC CHAIN 246 491
CC CHAIN 492 836
CC CHAIN 837 980
CC CHAIN 981 1087
CC CHAIN 1088 1422
CC CHAIN 1423 1496
CC CHAIN 1497 1519
CC CHAIN 1520 1738
CC CHAIN 1739 2227
CC VARIANT 77 77
CC VARIANT 764 764
CC VARIANT 821 821
CC VARIANT 1052 1052
CC VARIANT 1062 1062
CC VARIANT 1118 1118
CC VARIANT 1151 1151
CC VARIANT 1163 1163
CC VARIANT 1277 1277
CC VARIANT 1500 1500
CC VARIANT 1805 1805
CC VARIANT 1930 1930
CC SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;

Query Match 99.7%; Score 1056; DB 1; Length 2227;
Best Local Similarity 99.5%; Pred. No. 4.3e-87;
Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRREESHIECKRPYKELELVGKORLKYAQBELSENVLPPEPKKKGLFSQAK 60
Db 779 VDDPSEEDRREESHIECKRPYKELELVGKORLKYAQBELSENVLPPEPKKKGLFSQAK 838

Qy 61 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRINDEX 120
Db 839 ISLFYTEEHEIMKFSWGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRINDEX 898

Qy 121 WTEMODKIVSLIEKTSKWKYSKNVPHGMDLBEIAANSKDFPMSTDLCPFLHWIN 180
Db 899 WTEMODKIVSLIEKTSKWKYSKNVPHGMDLBEIAANSKDFPMSTDLCPFLHWIN 958

Db 181 PKKINLADRLGLSGVQEIKEQ 202
959 PKKINLADRLGLSGVQEIKEQ 980

RESULT 3

```

POLG_HPAV STANDARD; PRT: 2227 AA.
ID POLG_HPAV
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptide (Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)).
OS Hepatitis A virus (strain MBB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12100;
NM NCBI [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wiesel T., Kiehn R., Wimmer E.,
Deinhardt F.;
RA "The entire nucleotide sequence of the genome of human hepatitis A
virus (isolate MBB)."
Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PMT: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20773; AAA45474.1; -.
DR PIR; J50303; GNNYTH.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).
FT CHAIN 246 491 COAT PROTEIN VP2 (P1B).
FT CHAIN 492 836 COAT PROTEIN VP3 (P1C).
FT CHAIN 837 980 COAT PROTEIN VP1 (P1D).
FT CHAIN 981 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1496 CORE PROTEIN P2C.
FT CHAIN 1497 1519 PROBABLE PROTEIN P3A.
FT CHAIN 1520 1738 PROBABLE PROTEIN P3B.
FT CHAIN 1739 2227 PROBABLE PROTEIN P3C.
FT CHAIN 2227 2227 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2227 AA; 2511425 MW; EC983ED2A7C86349 CRC64;
Query Match 99.2%; Score 1050; DB 1; Length 2227;
Best Local Similarity 99.0%; Pred. No. 1.5e-86;
Matches 200; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

121 WTEKDDKIVSLIEKFTSNKYSKVPFPHGMIDLEEIAANSKDPFNMSETDLCFLHWIN 180
DB WTEKDDKIVSLIEKFTSNKYSKVPFPHGMIDLEEIAANSKDPFNMSETDLCFLHWIN 958
QY 181 PKKTLADRMGLSGVQRIKQ 202
DB 959 PKKTLADRMGLSGVQRIKQ 980
RESULT 4
POLG_HPAV STANDARD; PRT: 2226 AA.
ID POLG_HPAV
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptide (Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)).
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12095;
NM NCBI [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
Ramon S.M., Jansen R.W.,
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
variants arising during persistent infection: evidence for genetic
recombination."
J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PMT: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59809; AAA45469.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).
FT CHAIN 246 491 COAT PROTEIN VP2 (P1B).
FT CHAIN 492 794 COAT PROTEIN VP3 (P1C).
FT CHAIN 795 900 COAT PROTEIN VP1 (P1D).
FT CHAIN 901 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1495 CORE PROTEIN P2C.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3A.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3B.
FT CHAIN 1738 2226 PROBABLE PROTEIN P3C.
FT CHAIN 2226 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 2511107 MW; 403B4CA80B09BF75 CRC64;
Query Match 99.0%; Score 1048; DB 1; Length 2226;
Best Local Similarity 98.5%; Pred. No. 2.3e-86;
Matches 199; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 VDDPSEEDRPFRESHTECKRPYKELRLEVKGORLKYAOEELSENEVLPPPRKMKGLFSQAK 60
DB 779 VDDPSEEDRPFRESHTECKRPYKELRLEVKGORLKYAOEELSENEVLPPPRKMKGLFSQAN 838
QY 61 ISLFYTEEHEIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 120
DB 839 ISLFYTEEHEIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 898
QY 121 WTEMKDDKIVSLIEKFTSNKYWSKVPFHGMIDLEELIANSKDFPNNSETDLCFLLHMLN 180
DB 899 WTEMKDDKIVSLIEKFTSNKYWSKVPFHGMIDLEELIANSKDFPNNSETDLCFLLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

RESULT 5
POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
AT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OX NCBI_Taxid=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."
RT J. Virol. 65:2056-2065 (1991).
RL J. Virol. 65:2056-2065 (1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59808; AAA45467.1; -.
CC MEROPS: C03.005; -.
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR001205; RNA_pol_P3D.
CC Pfam: PF00680; RNA_dep_RNA_pol.1.
CC Pfam: PF00910; RNA_helicase.1.
CC PolyProtein: Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.
CC CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC CHAIN 492 794 COAT PROTEIN VP1 (PID).
CC CHAIN 795 900 CORE PROTEIN P2A.
CC CHAIN 901 1087 CORE PROTEIN P2B.
CC CHAIN 1088 1422 CORE PROTEIN P2C.
CC CHAIN 1423 1495 PROBABLE PROTEIN P3A.

```

```

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396CDB6 CRC64;

Query Match 99.0%; Score 1048; DB 1; Length 2226;
Best Local Similarity 98.5%; Pred. No. 2,36-86;
Matches 199; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDDPSEEDRPFRESHTECKRPYKELRLEVKGORLKYAOEELSENEVLPPPRKMKGLFSQAK 60
DB 779 VDDPSEEDRPFRESHTECKRPYKELRLEVKGORLKYAOEELSENEVLPPPRKMKGLFSQAN 838
QY 61 ISLFYTEEHEIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 120
DB 839 ISLFYTEEHEIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLINDEX 898
QY 121 WTEMKDDKIVSLIEKFTSNKYWSKVPFHGMIDLEELIANSKDFPNNSETDLCFLLHMLN 180
DB 899 WTEMKDDKIVSLIEKFTSNKYWSKVPFHGMIDLEELIANSKDFPNNSETDLCFLLHMLN 958
QY 181 PKKINLADRMGLSGVOEIKEQ 202
DB 959 PKKINLADRMGLSGVOEIKEQ 980

RESULT 6
POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
AT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OX NCBI_Taxid=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."
RT J. Virol. 65:2056-2065 (1991).
RL J. Virol. 65:2056-2065 (1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59810; AAA45468.1; -.
CC MEROPS: C03.005; -.
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR001205; RNA_pol_P3D.
CC Pfam: PF00680; RNA_dep_RNA_pol.1.
CC Pfam: PF00910; RNA_helicase.1.
CC PolyProtein: Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.

```

Med Apr 2 12:54:38 2003

```

FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CDB5A91D6B4E2BF CRC64;

Query Match 98.6%; Score 1044; DB 1; Length 2226;
Best Local Similarity 98.0%; Pred. No. 5.2e-86;
Matches 198; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRPFESHIECKRPYKELRLEVQKORLKYAOEELSNEVLPPRRKMKGLFSSQAK 60
DB 779 VDDPSEEDKPFESHI ECKRPYKELRLEVQKORLKYAOEELSNEVLPPRRKMKGLFSSQAN 838
61 ISLFTTEHEHEIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRILRLNDEK 120
DB 839 ISLFTTEHEHEMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRILRLNDEK 898
QY 121 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMLDLEEIANSKDPFNMSETDLCFLHMLN 180
DB 899 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMLDLEEIANSKDPFNMSETDLCFLHMLN 958
QY 181 PKINLADRMGLSGVQEIKEQ 202
DB 959 PKINLADRMGLSGVQEIKEQ 980

RESULT 7
POLG_HPAVC STANDARD; PRT; 2230 AA.
ID POLG_HPAVC
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
CC NCBI_TaxID=12102;
QX [1]
SEQUENCE FROM N.A.
RA MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Batayan M.S., Ticehurst J.R.,
RA Purcell R.H.,
RA "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains.";
RT J. Gen. Virol. 72:1677-1683(1991).
RL [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RX Batayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RX Sverdlov E.D., Chiznikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; D00924; BA00766.1; -.
DR EMBL; X15461; CA33490.1; -.
DR PIR; A30470; GNNVSA.
DR PIR; S04885; S04885.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27 COAT PROTEIN VP4 (PIA).
FT CHAIN 28 249 COAT PROTEIN VP2 (PIB).
FT CHAIN 250 495 COAT PROTEIN VP3 (PIC).
FT CHAIN 496 795 COAT PROTEIN VP1 (PID).
FT CHAIN 796 984 CORE PROTEIN P2A.
FT CHAIN 985 1091 CORE PROTEIN P2B.
FT CHAIN 1092 1426 CORE PROTEIN P2C.
FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2230 AA; 251236 MW; 87B3230E324E1F19 CRC64;

Query Match 91.7%; Score 971; DB 1; Length 2230;
Best Local Similarity 89.6%; Pred. No. 2e-79;
Matches 181; Conservative 15; Mismatches 6; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRPFESHIECKRPYKELRLEVQKORLKYAOEELSNEVLPPRRKMKGLFSSQAK 60
DB 783 VDDPSEADKPFESHI ECKRPYKELRLEVQKORLKYAOEELSNEVLPPRRKMKGLFSSQAK 842
QY 61 ISLFTTEHEHEIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRILRLNDEK 120
DB 843 ISLFTTEHEHEIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRILRLNDEK 902
QY 121 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMLDLEEIANSKDPFNMSETDLCFLHMLN 180
DB 903 WTEMKDKIVSLIEKFTSNKYKSKVNPFGMLDLEEIANSKDPFNMSETDLCFLHMLN 962
QY 181 PKINLADRMGLSGVQEIKEQ 202
DB 963 PKINLADRMGLSGVQEIKEQ 984

RESULT 8
POLG_HPAVC STANDARD; PRT; 852 AA.
ID POLG_HPAVC
AC P06442; O83741; O83742;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A (Fragment)].
DE P2A (Fragment)].
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
CC NCBI_TaxID=12097;
QX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185648; PubMed=2985793;
RX Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
RX Young A., Mitra S.W.,
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
RL J. Virol. 54:247-255(1985).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,

```

CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M10033; AAA45470.1; -
CC PIR: A03904; GNNYHA.
CC POLG_HPAVT; Coat protein; Core protein.
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
CC FT CHAIN 837 >852 CORE PROTEIN P2A.
CC FT NON TER 852 852
CC FT SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;
CC
CC Query Match 36.4%; Score 386; DB 1; Length 852;
CC Best Local Similarity 100.0%; Pred. No. 4.1e-27;
CC Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 VDDPSEDRRFRSHIECKRKYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKGLFSQAK 60
CC |||||
CC DB 779 VDDPSEDRRFRSHIECKRKYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKGLFSQAK 838
CC |||||
CC QY 61 ISLFYEEHEIMKE 74
CC |||||
CC DB 839 ISLFYEEHEIMKE 852
CC |||||
CC
CC RESULT 9
CC POLG_HPAVT STANDARD; PRT; 341 AA.
CC AC P13672;
CC DT 01-JAN-1990 (Rel. 13, Created)
CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Genome polyprotein [contains: Coat proteins VP1 TO VP3; Core protein
CC P2A] (Fragment).
CC OS Hepatitis A virus (strain LCD-1).
CC NCBI_TaxID=12093;
CC NCBI_TaxID=12093;
CC RP MEDLINE=89263805; PubMed=2542903;
CC RA Andonov A.P., Lau P., Chaudhary R.;
CC RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
CC hepatitis A virus (HAV).";
CC RT Nucleic Acids Res. 17:3594-3594(1989).
CC CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X14666; CAA32794.1; -
CC PIR: S04137; S04137.
CC KW Polyprotein; Coat protein; Core protein.

FT NON TER 1 1
FT CHAIN <1 1 COAT PROTEIN VP3 (IC).
FT CHAIN 2 340 COAT PROTEIN VP1 (ID).
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON TER 341 341
FT SEQUENCE 341 AA; 38003 MW; 066918289BF126F5 CRC64;
CC
CC Query Match 28.8%; Score 305; DB 1; Length 341;
CC Best Local Similarity 98.3%; Pred. No. 2.7e-20;
CC Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 VDDPSEDRRFRSHIECKRKYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKGLFSQAK 59
CC |||||
CC DB 283 VDDPSEDRRFRSHIECKRKYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKGLFSQAK 341
CC |||||
CC
CC RESULT 10
CC POLG_HPAVT STANDARD; PRT; 839 AA.
CC AC P31788;
CC DT 01-JUL-1993 (Rel. 26, Created)
CC DT 01-JUL-1993 (Rel. 26, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core protein
CC P2A] (Fragment).
CC OS Simian hepatitis A virus (strain CY-145).
CC NCBI_TaxID=31707;
CC NCBI_TaxID=31707;
CC RP MEDLINE=91311421; PubMed=1649902;
CC RA Naitan O.V., Margolis H.S., Robertson B.H., Brinton M.A.;
CC RT "Sequence analysis of a new hepatitis A virus naturally infecting
CC cynomolgus macaques (Macaca fascicularis).";
CC RT J. Gen. Virol. 72:1685-1688(1991).
CC CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59286; AAA45473.1; -
CC PIR: J01180; GNNYS2.
CC POLG_HPAVT; Coat protein; Core protein.
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
CC FT CHAIN ? >839 CORE PROTEIN P2A.
CC FT NON TER 839 839
CC FT SEQUENCE 839 AA; 93825 MW; 2CACC4BD1E192DBC CRC64;
CC
CC Query Match 26.2%; Score 277.5; DB 1; Length 839;
CC Best Local Similarity 85.5%; Pred. No. 2.5e-17;
CC Matches 53; Conservative 7; Mismatches 1; Indels 1; Gaps 1;
CC
CC QY 1 VDDPSEDRRFRSHIECKRKYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKGLFSQAK 60
CC |||||
CC DB 779 VDDPSEDRRFRSHIECKRKYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKGLFSQAK 837
CC |||||
CC QY 61 IS 62
CC |||||
CC DB 838 IS 839
CC |||||

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bernaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gwinnich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima U., Mazzaletti U., Momberts P.,
 RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohatsu S.,
 RA Hayashizaki Y.
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1- SIMILARITY: TO THE N-TERMINAL OF THREONYL-TRNA SYNTHETASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF239728; AAF4697.1; -
 DR EMBL; BC012274; AAH12274.1; -
 DR EMBL; BC016561; AAH16561.1; -
 DR EMBL; AK007681; BAB25185.1; -
 DR MGD; MGI:1351620; MRP139.
 DR InterPro; IPR004095; TGS_dom.
 DR Pfam; PF02824; TGS; 1.
 KM Ribosomal protein; Mitochondrion.
 FT CONFLICT 94 94 K -> R (IN REF. 2; AAH16561).
 FT CONFLICT 125 125 V -> L (IN REF. 2; AAH16561).
 FT CONFLICT 244 244 L -> F (IN REF. 1).
 FT CONFLICT 276 278 ERF -> GRS (IN REF. 2; AAH16561).
 FT CONFLICT 278 278 F -> S (IN REF. 2; AAH12274 AND 3).
 SQ SEQUENCE 297 AA; 34559 MW; 7E0461BCB875C57F CRC64;

Query Match 8.0%; Score 85; DB 1; Length 297;
 Best Local Similarity 27.8%; Pred. No. 1.7; Mismatches 37; Indels 56; Gaps 8;
 Matches 42; Conservative 14;

QY 73 KPSWGVATDTRAL--RRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKMTKDKIV 130
 DB 154 KENLRSEFTDAHALIVRDLPEF-----TLVDARV----- 183
 QY 131 SLIEKTSNKNYKKNF-----FHGMDLEIANGKDFPMNSE-----TDLCF-- 174
 DB 184 -ALEIFQHNKY--KVFIEBKASQNPRIVKLRIG--DFIDVSBGLPIPTSVCFQY 236
 QY 175 ---LHMLNPKKINLADRMGLSGVOEIKQ 202
 DB 237 EVSAVHNLPNQNLIRFGSLPPTHRAQ 267

RESULT 14
 RIR2 CAEEL STANDARD; PRT; 381 AA.
 AC P42170;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
 DE (Ribonucleoside reductase).
 GN RNR-2 OR C03C10.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OC NCB1_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Berks M.;
 RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
 CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
 CC thioresoxin + H(2)O = ribonucleoside diphosphate + reduced
 CC thioresoxin.
 CC -1- COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).
 CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
 CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
 CC SMALL CHAIN FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; Z35637; CA84688.1; -
 DR HSSP; P1157; IXSM.
 DR WormPep; C03C10.3; CE00874.
 DR InterPro; IPR000358; Ribonuc_red_sm; 1.
 DR Pfam; PF00268; Ribonuc_red_sm; 1.
 DR PROSITE; PS00368; RIBORED_SMALL; 1.
 KM Oxidoreductase; DNA replication; Iron.
 FT METAL 130 130 IRON 1 (BY SIMILARITY).
 FT METAL 161 161 IRON 1 AND 2 (BY SIMILARITY).
 FT METAL 164 164 IRON 1 (BY SIMILARITY).
 FT METAL 224 224 IRON 2 (BY SIMILARITY).
 FT METAL 258 258 IRON 2 (BY SIMILARITY).
 FT METAL 261 261 IRON 2 (BY SIMILARITY).
 FT ACT SITE 168 168 BY SIMILARITY.
 SQ SEQUENCE 381 AA; 44289 MW; 75497147ABF36C59 CRC64;

Query Match 8.0%; Score 85; DB 1; Length 381;
 Best Local Similarity 21.0%; Pred. No. 2.3; Mismatches 58; Indels 84; Gaps 12;
 Matches 48; Conservative 39;

QY 23 KEL-RLNEVGKQKLVKAAOELISNEVLPPPKMKGLPSQA-----KISLFTEHEIMK 73
 DB 30 KELEKLEIVDQKKAASAEETNNE-----SEVVELADDEMDLDDNRPVFLPKHHDITN 84
 QY 74 FSWRGVATDTRALRRFGSLAAGRSVWTL-EMDAGVLTGRLRLNDEKMTKDK- 128
 DB 85 FYKKAIVA-----SFWTVEVDLCK-----DMND--WEKNAGDEQYFI 119
 QY 129 -----IV-SLIEKTSNKNYKKNFPHG-MLDLEIANGKDFPMSETDL- 172
 DB 120 SRLAFPAASDGIIVNMLCERPSNEVOVSERFFGFOIAINI--HSEMYSKLIETIYR 177
 QY 173 -----CPLHMLNPKKINLADRMGLSGVOEI 199
 DB 178 DETERNTLFNAVDEEFYKKADWALRWISDKKASFAERLIFAAVEGI 226

RESULT 15
 CYA9_HUMAN STANDARD; PRT; 1353 AA.
 AC O60503; Q9UGP2; Q60273; Q9BWT4;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adenylate cyclase, type IX (EC 4.6.1.1) (ATP pyrophosphate-lyase)
 DE (Adenyl1 cyclase).
 GN ADCY9 OR KIAA0520.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCB1_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98292499; PubMed=9628827;

RA Hacker B.M., Tomlinson J.E., Wayman G.A., Sultana R., Chan G.,
RA Villacres E., Distche C., Storm D.R.;
RT "Cloning, chromosomal mapping, and regulatory properties of the human
RT type 9 adenylyl cyclase (ADCY9).";
RL Genomics 50:97-104(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Paterson J.M., Smith S.M., Simpson J., Grace O.C., Bell J.E.,
RA Antoni F.A.;
RT "Cloning and characterization of human adenylyl cyclase IX:
RT differential mRNA regulation and inhibition by Ca2+/calciueinin.";
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Toyota T., Yamada K., Meerbux J., Hattori E., Saito K.,
RA Yoshitaka K., Shimizu H., Nankai M., Toru M., Deterra-Wadleigh S.D.,
RA Yoshikawa T.;
RT "Mutation screening, case control study and transmission
RT disequilibrium analysis of adenylyl cyclase type 9 (ADCY9) gene in
RT functional psychoses.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 788-1353 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Ref. 5:31-39(1998).
CC -1- FUNCTION: May play a fundamental role in situations where fine
CC interplay between intracellular Ca(2+) and cAMP determines the
CC cellular function. May be a physiologically relevant docking site
CC for calcineurin (by similarity).
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -1- ENZYME REGULATION: Insensitive to Ca(2+)/calmodulin, forskolin and
CC bombesin. Stimulated by beta-adrenergic receptor activation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 1 FKBP-TYPE PPIASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF036927; AAC24201.1; -;
DR EMBL; AJ133123; CAB55084.1; -;
DR EMBL; AY028959; AAK29464.1; -;
DR EMBL; AY028949; AAK29464.1; JOINED.
DR EMBL; AY028950; AAK29464.1; JOINED.
DR EMBL; AY028951; AAK29464.1; JOINED.
DR EMBL; AY028952; AAK29464.1; JOINED.
DR EMBL; AY028953; AAK29464.1; JOINED.
DR EMBL; AY028954; AAK29464.1; JOINED.
DR EMBL; AY028955; AAK29464.1; JOINED.
DR EMBL; AY028956; AAK29464.1; JOINED.
DR EMBL; AY028957; AAK29464.1; JOINED.
DR EMBL; AB011092; BAA25446.1; -;
DR HSSP; P26769; IAB8.
DR Genew; HGNC:240; ADCY9.
DR MIM; 603302; -;
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 2.
DR SMART; SM00044; CYCC; 2_cyc; 2.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 2.

KM Lyase; cAMP synthesis; Transmembrane; Glycoprotein; Repeat;
KW Isomerase; Rotamase.
FT DOMAIN 1 117
FT TRANSMEM 118 138
FT DOMAIN 139 141
FT TRANSMEM 142 162
FT DOMAIN 163 171
FT TRANSMEM 172 192
FT DOMAIN 193 215
FT TRANSMEM 216 235
FT DOMAIN 236 241
FT TRANSMEM 242 259
FT DOMAIN 260 280
FT TRANSMEM 281 301
FT DOMAIN 302 321
FT TRANSMEM 322 341
FT DOMAIN 342 361
FT TRANSMEM 362 381
FT DOMAIN 382 401
FT TRANSMEM 402 421
FT DOMAIN 422 441
FT TRANSMEM 442 461
FT DOMAIN 462 481
FT TRANSMEM 482 501
FT DOMAIN 502 521
FT TRANSMEM 522 541
FT DOMAIN 542 561
FT TRANSMEM 562 581
FT DOMAIN 582 601
FT TRANSMEM 602 621
FT DOMAIN 622 641
FT TRANSMEM 642 661
FT DOMAIN 662 681
FT TRANSMEM 682 701
FT DOMAIN 702 721
FT TRANSMEM 722 741
FT DOMAIN 742 761
FT TRANSMEM 762 781
FT DOMAIN 782 801
FT TRANSMEM 802 821
FT DOMAIN 822 841
FT TRANSMEM 842 861
FT DOMAIN 862 881
FT TRANSMEM 882 901
FT DOMAIN 902 921
FT TRANSMEM 922 941
FT DOMAIN 942 961
FT TRANSMEM 962 981
FT DOMAIN 982 1001
FT TRANSMEM 1002 1021
FT DOMAIN 1022 1041
FT TRANSMEM 1042 1061
FT DOMAIN 1062 1081
FT TRANSMEM 1082 1101
FT DOMAIN 1102 1121
FT TRANSMEM 1122 1141
FT DOMAIN 1142 1161
FT TRANSMEM 1162 1181
FT DOMAIN 1182 1201
FT TRANSMEM 1202 1221
FT DOMAIN 1222 1241
FT TRANSMEM 1242 1261
FT DOMAIN 1262 1281
FT TRANSMEM 1282 1301
FT DOMAIN 1302 1321
FT TRANSMEM 1322 1341
FT DOMAIN 1342 1361
FT TRANSMEM 1362 1381
FT DOMAIN 1382 1401
FT TRANSMEM 1402 1421
FT DOMAIN 1422 1441
FT TRANSMEM 1442 1461
FT DOMAIN 1462 1481
FT TRANSMEM 1482 1501
FT DOMAIN 1502 1521
FT TRANSMEM 1522 1541
FT DOMAIN 1542 1561
FT TRANSMEM 1562 1581
FT DOMAIN 1582 1601
FT TRANSMEM 1602 1621
FT DOMAIN 1622 1641
FT TRANSMEM 1642 1661
FT DOMAIN 1662 1681
FT TRANSMEM 1682 1701
FT DOMAIN 1702 1721
FT TRANSMEM 1722 1741
FT DOMAIN 1742 1761
FT TRANSMEM 1762 1781
FT DOMAIN 1782 1801
FT TRANSMEM 1802 1821
FT DOMAIN 1822 1841
FT TRANSMEM 1842 1861
FT DOMAIN 1862 1881
FT TRANSMEM 1882 1901
FT DOMAIN 1902 1921
FT TRANSMEM 1922 1941
FT DOMAIN 1942 1961
FT TRANSMEM 1962 1981
FT DOMAIN 1982 2001
FT TRANSMEM 2002 2021
FT DOMAIN 2022 2041
FT TRANSMEM 2042 2061
FT DOMAIN 2062 2081
FT TRANSMEM 2082 2101
FT DOMAIN 2102 2121
FT TRANSMEM 2122 2141
FT DOMAIN 2142 2161
FT TRANSMEM 2162 2181
FT DOMAIN 2182 2201
FT TRANSMEM 2202 2221
FT DOMAIN 2222 2241
FT TRANSMEM 2242 2261
FT DOMAIN 2262 2281
FT TRANSMEM 2282 2301
FT DOMAIN 2302 2321
FT TRANSMEM 2322 2341
FT DOMAIN 2342 2361
FT TRANSMEM 2362 2381
FT DOMAIN 2382 2401
FT TRANSMEM 2402 2421
FT DOMAIN 2422 2441
FT TRANSMEM 2442 2461
FT DOMAIN 2462 2481
FT TRANSMEM 2482 2501
FT DOMAIN 2502 2521
FT TRANSMEM 2522 2541
FT DOMAIN 2542 2561
FT TRANSMEM 2562 2581
FT DOMAIN 2582 2601
FT TRANSMEM 2602 2621
FT DOMAIN 2622 2641
FT TRANSMEM 2642 2661
FT DOMAIN 2662 2681
FT TRANSMEM 2682 2701
FT DOMAIN 2702 2721
FT TRANSMEM 2722 2741
FT DOMAIN 2742 2761
FT TRANSMEM 2762 2781
FT DOMAIN 2782 2801
FT TRANSMEM 2802 2821
FT DOMAIN 2822 2841
FT TRANSMEM 2842 2861
FT DOMAIN 2862 2881
FT TRANSMEM 2882 2901
FT DOMAIN 2902 2921
FT TRANSMEM 2922 2941
FT DOMAIN 2942 2961
FT TRANSMEM 2962 2981
FT DOMAIN 2982 3001
FT TRANSMEM 3002 3021
FT DOMAIN 3022 3041
FT TRANSMEM 3042 3061
FT DOMAIN 3062 3081
FT TRANSMEM 3082 3101
FT DOMAIN 3102 3121
FT TRANSMEM 3122 3141
FT DOMAIN 3142 3161
FT TRANSMEM 3162 3181
FT DOMAIN 3182 3201
FT TRANSMEM 3202 3221
FT DOMAIN 3222 3241
FT TRANSMEM 3242 3261
FT DOMAIN 3262 3281
FT TRANSMEM 3282 3301
FT DOMAIN 3302 3321
FT TRANSMEM 3322 3341
FT DOMAIN 3342 3361
FT TRANSMEM 3362 3381
FT DOMAIN 3382 3401
FT TRANSMEM 3402 3421
FT DOMAIN 3422 3441
FT TRANSMEM 3442 3461
FT DOMAIN 3462 3481
FT TRANSMEM 3482 3501
FT DOMAIN 3502 3521
FT TRANSMEM 3522 3541
FT DOMAIN 3542 3561
FT TRANSMEM 3562 3581
FT DOMAIN 3582 3601
FT TRANSMEM 3602 3621
FT DOMAIN 3622 3641
FT TRANSMEM 3642 3661
FT DOMAIN 3662 3681
FT TRANSMEM 3682 3701
FT DOMAIN 3702 3721
FT TRANSMEM 3722 3741
FT DOMAIN 3742 3761
FT TRANSMEM 3762 3781
FT DOMAIN 3782 3801
FT TRANSMEM 3802 3821
FT DOMAIN 3822 3841
FT TRANSMEM 3842 3861
FT DOMAIN 3862 3881
FT TRANSMEM 3882 3901
FT DOMAIN 3902 3921
FT TRANSMEM 3922 3941
FT DOMAIN 3942 3961
FT TRANSMEM 3962 3981
FT DOMAIN 3982 4001
FT TRANSMEM 4002 4021
FT DOMAIN 4022 4041
FT TRANSMEM 4042 4061
FT DOMAIN 4062 4081
FT TRANSMEM 4082 4101
FT DOMAIN 4102 4121
FT TRANSMEM 4122 4141
FT DOMAIN 4142 4161
FT TRANSMEM 4162 4181
FT DOMAIN 4182 4201
FT TRANSMEM 4202 4221
FT DOMAIN 4222 4241
FT TRANSMEM 4242 4261
FT DOMAIN 4262 4281
FT TRANSMEM 4282 4301
FT DOMAIN 4302 4321
FT TRANSMEM 4322 4341
FT DOMAIN 4342 4361
FT TRANSMEM 4362 4381
FT DOMAIN 4382 4401
FT TRANSMEM 4402 4421
FT DOMAIN 4422 4441
FT TRANSMEM 4442 4461
FT DOMAIN 4462 4481
FT TRANSMEM 4482 4501
FT DOMAIN 4502 4521
FT TRANSMEM 4522 4541
FT DOMAIN 4542 4561
FT TRANSMEM 4562 4581
FT DOMAIN 4582 4601
FT TRANSMEM 4602 4621
FT DOMAIN 4622 4641
FT TRANSMEM 4642 4661
FT DOMAIN 4662 4681
FT TRANSMEM 4682 4701
FT DOMAIN 4702 4721
FT TRANSMEM 4722 4741
FT DOMAIN 4742 4761
FT TRANSMEM 4762 4781
FT DOMAIN 4782 4801
FT TRANSMEM 4802 4821
FT DOMAIN 4822 4841
FT TRANSMEM 4842 4861
FT DOMAIN 4862 4881
FT TRANSMEM 4882 4901
FT DOMAIN 4902 4921
FT TRANSMEM 4922 4941
FT DOMAIN 4942 4961
FT TRANSMEM 4962 4981
FT DOMAIN 4982 5001
FT TRANSMEM 5002 5021
FT DOMAIN 5022 5041
FT TRANSMEM 5042 5061
FT DOMAIN 5062 5081
FT TRANSMEM 5082 5101
FT DOMAIN 5102 5121
FT TRANSMEM 5122 5141
FT DOMAIN 5142 5161
FT TRANSMEM 5162 5181
FT DOMAIN 5182 5201
FT TRANSMEM 5202 5221
FT DOMAIN 5222 5241
FT TRANSMEM 5242 5261
FT DOMAIN 5262 5281
FT TRANSMEM 5282 5301
FT DOMAIN 5302 5321
FT TRANSMEM 5322 5341
FT DOMAIN 5342 5361
FT TRANSMEM 5362 5381
FT DOMAIN 5382 5401
FT TRANSMEM 5402 5421
FT DOMAIN 5422 5441
FT TRANSMEM 5442 5461
FT DOMAIN 5462 5481
FT TRANSMEM 5482 5501
FT DOMAIN 5502 5521
FT TRANSMEM 5522 5541
FT DOMAIN 5542 5561
FT TRANSMEM 5562 5581
FT DOMAIN 5582 5601
FT TRANSMEM 5602 5621
FT DOMAIN 5622 5641
FT TRANSMEM 5642 5661
FT DOMAIN 5662 5681
FT TRANSMEM 5682 5701
FT DOMAIN 5702 5721
FT TRANSMEM 5722 5741
FT DOMAIN 5742 5761
FT TRANSMEM 5762 5781
FT DOMAIN 5782 5801
FT TRANSMEM 5802 5821
FT DOMAIN 5822 5841
FT TRANSMEM 5842 5861
FT DOMAIN 5862 5881
FT TRANSMEM 5882 5901
FT DOMAIN 5902 5921
FT TRANSMEM 5922 5941
FT DOMAIN 5942 5961
FT TRANSMEM 5962 5981
FT DOMAIN 5982 6001
FT TRANSMEM 6002 6021
FT DOMAIN 6022 6041
FT TRANSMEM 6042 6061
FT DOMAIN 6062 6081
FT TRANSMEM 6082 6101
FT DOMAIN 6102 6121
FT TRANSMEM 6122 6141
FT DOMAIN 6142 6161
FT TRANSMEM 6162 6181
FT DOMAIN 6182 6201
FT TRANSMEM 6202 6221
FT DOMAIN 6222 6241
FT TRANSMEM 6242 6261
FT DOMAIN 6262 6281
FT TRANSMEM 6282 6301
FT DOMAIN 6302 6321
FT TRANSMEM 6322 6341
FT DOMAIN 6342 6361
FT TRANSMEM 6362 6381
FT DOMAIN 6382 6401
FT TRANSMEM 6402 6421
FT DOMAIN 6422 6441
FT TRANSMEM 6442 6461
FT DOMAIN 6462 6481
FT TRANSMEM 6482 6501
FT DOMAIN 6502 6521
FT TRANSMEM 6522 6541
FT DOMAIN 6542 6561
FT TRANSMEM 6562 6581
FT DOMAIN 6582 6601
FT TRANSMEM 6602 6621
FT DOMAIN 6622 6641
FT TRANSMEM 6642 6661
FT DOMAIN 6662 6681
FT TRANSMEM 6682 6701
FT DOMAIN 6702 6721
FT TRANSMEM 6722 6741
FT DOMAIN 6742 6761
FT TRANSMEM 6762 6781
FT DOMAIN 6782 6801
FT TRANSMEM 6802 6821
FT DOMAIN 6822 6841
FT TRANSMEM 6842 6861
FT DOMAIN 6862 6881
FT TRANSMEM 6882 6901
FT DOMAIN 6902 6921
FT TRANSMEM 6922 6941
FT DOMAIN 6942 6961
FT TRANSMEM 6962 6981
FT DOMAIN 6982 7001
FT TRANSMEM 7002 7021
FT DOMAIN 7022 7041
FT TRANSMEM 7042 7061
FT DOMAIN 7062 7081
FT TRANSMEM 7082 7101
FT DOMAIN 7102 7121
FT TRANSMEM 7122 7141
FT DOMAIN 7142 7161
FT TRANSMEM 7162 7181
FT DOMAIN 7182 7201
FT TRANSMEM 7202 7221
FT DOMAIN 7222 7241
FT TRANSMEM 7242 7261
FT DOMAIN 7262 7281
FT TRANSMEM 7282 7301
FT DOMAIN 7302 7321
FT TRANSMEM 7322 7341
FT DOMAIN 7342 7361
FT TRANSMEM 7362 7381
FT DOMAIN 7382 7401
FT TRANSMEM 7402 7421
FT DOMAIN 7422 7441
FT TRANSMEM 7442 7461
FT DOMAIN 7462 7481
FT TRANSMEM 7482 7501
FT DOMAIN 7502 7521
FT TRANSMEM 7522 7541
FT DOMAIN 7542 7561
FT TRANSMEM 7562 7581
FT DOMAIN 7582 7601
FT TRANSMEM 7602 7621
FT DOMAIN 7622 7641
FT TRANSMEM 7642 7661
FT DOMAIN 7662 7681
FT TRANSMEM 7682 7701
FT DOMAIN 7702 7721
FT TRANSMEM 7722 7741
FT DOMAIN 7742 7761
FT TRANSMEM 7762 7781
FT DOMAIN 7782 7801
FT TRANSMEM 7802 7821
FT DOMAIN 7822 7841
FT TRANSMEM 7842 7861
FT DOMAIN 7862 7881
FT TRANSMEM 7882 7901
FT DOMAIN 7902 7921
FT TRANSMEM 7922 7941
FT DOMAIN 7942 7961
FT TRANSMEM 7962 7981
FT DOMAIN 7982 8001
FT TRANSMEM 8002 8021
FT DOMAIN 8022 8041
FT TRANSMEM 8042 8061
FT DOMAIN 8062 8081
FT TRANSMEM 8082 8101
FT DOMAIN 8102 8121
FT TRANSMEM 8122 8141
FT DOMAIN 8142 8161
FT TRANSMEM 8162 8181
FT DOMAIN 8182 8201
FT TRANSMEM 8202 8221
FT DOMAIN 8222 8241
FT TRANSMEM 8242 8261
FT DOMAIN 8262 8281
FT TRANSMEM 8282 8301
FT DOMAIN 8302 8321
FT TRANSMEM 8322 8341
FT DOMAIN 8342 8361
FT TRANSMEM 8362 8381
FT DOMAIN 8382 8401
FT TRANSMEM 8402 8421
FT DOMAIN 8422 8441
FT TRANSMEM 8442 8461
FT DOMAIN 8462 8481
FT TRANSMEM 8482 8501
FT DOMAIN 8502 8521
FT TRANSMEM 8522 8541
FT DOMAIN 8542 8561
FT TRANSMEM 8562 8581
FT DOMAIN 8582 8601
FT TRANSMEM 8602 8621
FT DOMAIN 8622 8641
FT TRANSMEM 8642 8661
FT DOMAIN 8662 8681
FT TRANSMEM 8682 8701
FT DOMAIN 8702 8721
FT TRANSMEM 8722 8741
FT DOMAIN 8742 8761
FT TRANSMEM 8762 8781
FT DOMAIN 8782 8801
FT TRANSMEM 8802 8821
FT DOMAIN 8822 8841
FT TRANSMEM 8842 8861
FT DOMAIN 8862 8881
FT TRANSMEM 8882 8901
FT DOMAIN 8902 8921
FT TRANSMEM 8922 8941
FT DOMAIN 8942 8961
FT TRANSMEM 8962 8981
FT DOMAIN 8982 9001
FT TRANSMEM 9002 9021
FT DOMAIN 9022 9041
FT TRANSMEM 9042 9061
FT DOMAIN 9062 9081
FT TRANSMEM 9082 9101
FT DOMAIN 9102 9121
FT TRANSMEM 9122 9141
FT DOMAIN 9142 9161
FT TRANSMEM 9162 9181
FT DOMAIN 9182 9201
FT TRANSMEM 9202 9221
FT DOMAIN 9222 9241
FT TRANSMEM 9242 9261
FT DOMAIN 9262 9281
FT TRANSMEM 9282 9301
FT DOMAIN 9302 9321
FT TRANSMEM 9322 9341
FT DOMAIN 9342 9361
FT TRANSMEM 9362 9381
FT DOMAIN 9382 9401
FT TRANSMEM 9402 9421
FT DOMAIN 9422 9441
FT TRANSMEM 9442 9461
FT DOMAIN 9462 9481
FT TRANSMEM 9482 9501
FT DOMAIN 9502 9521
FT TRANSMEM 9522 9541
FT DOMAIN 9542 9561
FT TRANSMEM 9562 9581
FT DOMAIN 9582 9601
FT TRANSMEM 9602 9621
FT DOMAIN 9622 9641
FT TRANSMEM 9642 9661
FT DOMAIN 9662 9681
FT TRANSMEM 9682 9701
FT DOMAIN 9702 9721
FT TRANSMEM 9722 9741
FT DOMAIN 9742 9761
FT TRANSMEM 9762 9781
FT DOMAIN 9782 9801
FT TRANSMEM 9802 9821
FT DOMAIN 9822 9841
FT TRANSMEM 9842 9861
FT DOMAIN 9862 9881
FT TRANSMEM 9882 9901
FT DOMAIN 9902 9921
FT TRANSMEM 9922 9941
FT DOMAIN 9942 9961
FT TRANSMEM 9962 9981
FT DOMAIN 9982 10001
FT TRANSMEM 10002 10021
FT DOMAIN 10022 10041
FT TRANSMEM 10042 10061
FT DOMAIN 10062 10081
FT TRANSMEM 10082 10101
FT DOMAIN 10102 10121
FT TRANSMEM 10122 10141
FT DOMAIN 10142 10161
FT TRANSMEM 10162 10181
FT DOMAIN 10182 10201
FT TRANSMEM 10202 10221
FT DOMAIN 10222 10241
FT TRANSMEM 10242 10261
FT DOMAIN 10262 10281
FT TRANSMEM 10282 10301
FT DOMAIN 10302 10321
FT TRANSMEM 10322 10341
FT DOMAIN 10342 10361
FT TRANSMEM 10362 10381
FT DOMAIN 10382 10401
FT TRANSMEM 10402 10421
FT DOMAIN 10422 10441
FT TRANSMEM 10442 10461
FT DOMAIN 10462 10481
FT TRANSMEM 10482 10501
FT DOMAIN 10502 10521
FT TRANSMEM 10522 10541
FT DOMAIN 10542 10561
FT TRANSMEM 10562 10581
FT DOMAIN 10582 10601
FT TRANSMEM 10602 10621
FT DOMAIN 10622 10641
FT TRANSMEM 10642 10661
FT DOMAIN 10662 10681
FT TRANSMEM 10682 10701
FT DOMAIN 10702 10721
FT TRANSMEM 10722 10741
FT DOMAIN 10742 10761
FT TRANSMEM 10762 10781
FT DOMAIN 10782 10801
FT TRANSMEM 10802 10821
FT DOMAIN 10822 10841
FT TRANSMEM 10842 10861
FT DOMAIN 10862 10881
FT TRANSMEM 10882 10901
FT DOMAIN 10902 10921
FT TRANSMEM 10922 10941
FT DOMAIN 10942 10961
FT TRANSMEM 10962 10981
FT DOMAIN 10982 11001
FT TRANSMEM 11002 11021
FT DOMAIN 11022 11041
FT TRANSMEM 11042 11061
FT DOMAIN 11062 11081
FT TRANSMEM 11082 11101
FT DOMAIN 11102 11121
FT TRANSMEM 11122 11141
FT DOMAIN 11142 11161
FT TRANSMEM 11162 11181
FT DOMAIN 11182 11201
FT TRANSMEM 11202 11221
FT DOMAIN 11222 11241
FT TRANSMEM 11242 11261
FT DOMAIN 11262 11281
FT TRANSMEM 11282 11301
FT DOMAIN 11302 11321
FT TRANSMEM 11322 11341
FT DOMAIN 11342 11361
FT TRANSMEM 11362 11381
FT DOMAIN 11382 11401
FT TRANSMEM 11402 11421
FT DOMAIN 11422 11441
FT TRANSMEM 11442 11461
FT DOMAIN 11462 11481
FT TRANSMEM 11482 11501
FT DOMAIN 11502 11521
FT TRANSMEM 11522 11541
FT DOMAIN 11542 11561
FT TRANSMEM 11562 11581
FT DOMAIN 11582 11601
FT TRANSMEM 11602 11621
FT DOMAIN 11622 11641
FT TRANSMEM 11642 11661
FT DOMAIN 11662 11681
FT TRANSMEM 11682 11701
FT DOMAIN 11702 11721
FT TRANSMEM 11722 11741
FT DOMAIN 11742 11761
FT TRANSMEM 11762 11781
FT DOMAIN 11782 11801
FT TRANSMEM 11802 11821
FT DOMAIN 11822 11841
FT TRANSMEM 11842 11861
FT DOMAIN 11862 11881
FT TRANSMEM 11882 11901
FT DOMAIN 11902 11921
FT TRANSMEM 11922 11941
FT DOMAIN 11942 11961
FT TRANSMEM 11962 11981
FT DOMAIN 11982 12001
FT TRANSMEM 12002 12021
FT DOMAIN 12022 12041
FT TRANSMEM 12042 12061
FT DOMAIN 12062 12081
FT TRANSMEM 12082 12101
FT DOMAIN 12102 12121
FT TRANSMEM 12122 12141
FT DOMAIN 12142 12161
FT TRANSMEM 12162 12181
FT DOMAIN 12182 12201
FT TRANSMEM 12202 12221
FT DOMAIN 12222 12241
FT TRANSMEM 12242 12261
FT DOMAIN 12262 12281
FT TRANSMEM 12282 12301
FT DOMAIN 12302 12321
FT TRANSMEM 12322 12341
FT DOMAIN 12342 12361
FT TRANSMEM 12362 12381
FT DOMAIN 12382 12401
FT TRANSMEM 12402 12421
FT DOMAIN 12422 12441
FT TRANSMEM 12442 12461
FT DOMAIN 12462 12481
FT TRANSMEM 12482 12501
FT DOMAIN 12502 12521
FT TRANSMEM 12522 12541
FT DOMAIN 12542 12561
FT TRANSMEM 12562 12581
FT DOMAIN 12582 12601
FT TRANSMEM 12602 12621
FT DOMAIN 12622 12641
FT TRANSMEM 12642 12661
FT DOMAIN 12662 12681
FT TRANSMEM 12682 12701
FT DOMAIN 12702 12721
FT TRANSMEM 12722 12741
FT DOMAIN 12742 12761
FT TRANSMEM 12762 12781
FT DOMAIN 12782 12801
FT TRANSMEM 12802 12821
FT DOMAIN 12822 12841
FT TRANSMEM 12842 12861
FT DOMAIN 12862 12881
FT TRANSMEM 12882 12901
FT DOMAIN 12902 12921
FT TRANSMEM 12922 12941
FT DOMAIN 12942 12961
FT TRANSMEM 12962 12981
FT DOMAIN 12982 13001
FT TRANSMEM 13002 13021
FT DOMAIN 13022 13041
FT TRANSMEM 13042 13061
FT DOMAIN 13062 13081
FT TRANSMEM 13082 13101
FT DOMAIN 13102 13121
FT TRANSMEM 13122 13141
FT DOMAIN 13142 13161
FT TRANSMEM 13162 13181
FT DOMAIN 13182 13201
FT TRANSMEM 13202 13221
FT DOMAIN 13222 13241
FT TRANSMEM 13242 13261
FT DOMAIN 13262 13281
FT TRANSMEM 13282 13301
FT DOMAIN 13302 13321
FT TRANSMEM 13322 13341
FT DOMAIN 13342 13361
FT TRANSMEM 13362 13381
FT DOMAIN 13382 13401
FT TRANSMEM 13402 13421
FT DOMAIN 13422 13441
FT TRANSMEM 13442 13461
FT DOMAIN 13462 13481
FT TRANSMEM 13482 13501
FT DOMAIN 13502 13521
FT TRANSMEM 13522 13541
FT DOMAIN 13542 13561
FT TRANSMEM 13562 13581
FT DOMAIN 13582 13601
FT TRANSMEM 13602 13621
FT DOMAIN 13622 13641
FT TRANSMEM 13642 13661
FT DOMAIN 13662 13681
FT TRANSMEM 13682 13701
FT DOMAIN 13702 13721
FT TRANSMEM 13722 13741
FT DOMAIN 13742 13761
FT TRANSMEM 13762 13781
FT DOMAIN 13782 13801
FT TRANSMEM 13802 13821
FT DOMAIN 13822 13841
FT TRANSMEM 13842 13861
FT DOMAIN 13862 13881
FT TRANSMEM 13882 13901
FT DOMAIN 13902 13921
FT TRANSMEM

THIS PAGE BLANK (USPTO)

QY 121 WTEMKDKIVSLIEKFTSNKYKSVNFPHGMLDLEIIAANSKDFPNNSETDLCFLHMLN 180
 DB 856 WTEMKDKIVSLIEKFTSNKYKSVNFPHGMLDLEIIAANSKDFPNNSETDLCFLHMLN 915
 QY 181 PKKINLADRLGLSGVQEIKEQ 202
 DB 916 PKKINLADRLGLSGVQEIKEQ 937

RESULT 2

Q05794 PRELIMINARY: PRT; 1161 AA.
 AC Q05794; Q67800; Q67801; Q67802; Q67803; Q67804; Q67805; Q67806;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Genome polyprotein (coat proteins VP1 to VP4; core proteins P2a to P2c; probable proteins P3a to P3c; RNA-directed RNA polymerase P3d (EC 2.7.7.48)) (Fragment).
 DE Hepatitis A virus.
 DE Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
 NC NCB1_TaxID=12092;
 OK NCB1_TaxID=12092;
 RN
 RP SEQUENCE FROM N.A.
 RA Sverdlov S.D., Tsarev S.A., Markova S.V., Vasilenko S.K.,
 RA Chirkov V.E., Petrov N.A., Kusov Y.Y., Nastashenko T.A.,
 RA Balayan M.S.,
 RA Mol. Gen. Microbiol. Virol. 6:129-133(1987).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC DR EMBL: X15464; CAA33492.1; -
 DR InterPro: IPR000408; Reg. chr. condens.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
 KW Hydroxylase; Thiol protease.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 >1161 CORE PROTEIN P2C.
 FT NON TER 1161 1161
 SQ SEQUENCE 1161 AA; 131131 MM; 38BB93789FEC3400 CRC64;

Query Match 100.0%; Score 1059; DB 12; Length 1161;
 Best Local Similarity 100.0%; Pred. No. 3, 6e-90;
 Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFESHIECKRPYKELRLEVQKORLKYAQBELSNEVLPKPKMGGLFSQAK 60
 DB 773 VDDPSEEDRRFESHIECKRPYKELRLEVQKORLKYAQBELSNEVLPKPKMGGLFSQAK 832
 QY 61 ISLFYTEHEHIMKFSWGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 833 ISLFYTEHEHIMKFSWGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 892
 QY 121 WTEMKDKIVSLIEKFTSNKYKSVNFPHGMLDLEIIAANSKDFPNNSETDLCFLHMLN 180
 DB 893 WTEMKDKIVSLIEKFTSNKYKSVNFPHGMLDLEIIAANSKDFPNNSETDLCFLHMLN 952
 QY 181 PKKINLADRLGLSGVQEIKEQ 202
 DB 953 PKKINLADRLGLSGVQEIKEQ 974

RESULT 3
 ID Q9DLJ2 PRELIMINARY: PRT; 2225 AA.
 AC Q9DLJ2;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Polyprotein.
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OK NCB1_TaxID=12092;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=L-A-1;
 RA Wang P.F., Jiang C.L., Liu J.Y., Zhang H.Y.;
 RA Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF314208; AAG5423.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR000408; Reg chr. condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00910; RNA_helicase; 1.
 DR PRINTS: PR00918; CALICIVIRUSN.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 SQ SEQUENCE 2225 AA; 251297 MM; EBACE41B043E5E9B CRC64;

Query Match 99.7%; Score 1056; DB 12; Length 2225;
 Best Local Similarity 99.5%; Pred. No. 1, 6e-89;
 Matches 201; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFESHIECKRPYKELRLEVQKORLKYAQBELSNEVLPKPKMGGLFSQAK 60
 DB 773 VDDPSEEDRRFESHIECKRPYKELRLEVQKORLKYAQBELSNEVLPKPKMGGLFSQAK 838
 QY 61 ISLFYTEHEHIMKFSWGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 833 ISLFYTEHEHIMKFSWGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
 QY 121 WTEMKDKIVSLIEKFTSNKYKSVNFPHGMLDLEIIAANSKDFPNNSETDLCFLHMLN 180
 DB 899 WTEMKDKIVSLIEKFTSNKYKSVNFPHGMLDLEIIAANSKDFPNNSETDLCFLHMLN 958
 QY 181 PKKINLADRLGLSGVQEIKEQ 202
 DB 959 PKKINLADRLGLSGVQEIKEQ 980

RESULT 4

Q9MMA2 PRELIMINARY: PRT; 2216 AA.
 ID Q9MMA2;
 AC Q9MMA2;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Polyprotein.
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OK NCB1_TaxID=12092;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=AH3;
 RC MEDLINE=21386014; PubMed=11495028;
 RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saisho H., Omata M.;
 RA "Analysis of full-length hepatitis A virus genome in sera from
 RT patients with fulminant and self-limited acute type A hepatitis";
 RL J. Hepatol. 35:112-119(2001).
 DR EMBL: AB020566; BAA35104.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR000408; Reg chr. condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00910; RNA_helicase; 1.

DR PRINTS: PR00918; CALICVIRUS.
DR PROSITE; PS00626; RCCL_2; UNKNOWN 1.
SQ SEQUENCE 2216 AA; 250209 MW; 1A9D93FEC21FBE82 CRC64;

Query Match 99.4%; Score 1053; DB 12; Length 2216;
Best Local Similarity 98.5%; Pred. No. 3e-89; Indels 0; Gaps 0;
Matches 199; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECRKPYKEIRLEVGKORLKYAOEELSNEVLP PPRKKKGLFSQAK 60
DB 779 VDDPSEEDRRFRESHIECRKPYKEIRLEVGKORLKYAOEELSNEVLP PPRKKKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEEHEIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDDKIVSLIEKFTSNKYSKYNPPHGMDLDEIANSKDFPMSETDLCFLLHMLN 180
DB 899 WTEMKDDKIVSLIEKFTSNKYSKYNPPHGMDLDEIANSKDFPMSETDLCFLLHMLN 958
DB 181 PKKINLADRMGLSGVQEIKEQ 202
DB 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 5
O67824 PRELIMINARY; PRT; 2218 AA.
AC O67824;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE GBM/PRBK RNA.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GBM;
RX MEDLINE=94076453; PubMed=8254770;
RA Graff J., Norman A., Feinstein S.M., Flehmig B.;
RT "Nucleotide sequence of wild-type hepatitis A virus GBM in comparison
to two cell culture adapted variants.";
RL J. Virol. 68:548-554(1994).
DR EMBL; X75214; CAAS3024.1; -
DR InterPro; IPR000408; Reg_chir_condens.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
PFam; PF00680; RNA_dep_RNA_pol; 1.
Pfam; PF00910; RNA_helicase; 1.
DR PROSITE; PS00626; RCCL_2; UNKNOWN 1.
FT CHAIN 1 785 P1 STRUCTURAL PROTEIN.
FT CHAIN 792 1422 P2 NONSTRUCTURAL PROTEIN.
FT CHAIN 1417 2218 P3 NONSTRUCTURAL PROTEIN.
SQ SEQUENCE 2218 AA; 250502 MW; CA72DF0922104C0E CRC64;

Query Match 99.4%; Score 1053; DB 12; Length 2218;
Best Local Similarity 98.5%; Pred. No. 3e-89; Indels 0; Gaps 0;
Matches 199; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECRKPYKEIRLEVGKORLKYAOEELSNEVLP PPRKKKGLFSQAK 60
DB 779 VDDPSEEDRRFRESHIECRKPYKEIRLEVGKORLKYAOEELSNEVLP PPRKKKGLFSQAK 832
QY 61 ISLFYTEEHEIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 833 ISLFYTEEHEIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 892
QY 121 WTEMKDDKIVSLIEKFTSNKYSKYNPPHGMDLDEIANSKDFPMSETDLCFLLHMLN 180
DB 893 WTEMKDDKIVSLIEKFTSNKYSKYNPPHGMDLDEIANSKDFPMSETDLCFLLHMLN 952

QY 181 PKKINLADRMGLSGVQEIKEQ 202
DB 953 PKKINLADRMGLSGVQEIKEQ 974

Query Match 99.4%; Score 1053; DB 12; Length 2227;
Best Local Similarity 98.5%; Pred. No. 3e-89; Indels 0; Gaps 0;
Matches 199; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
O67825 PRELIMINARY; PRT; 2227 AA.
AC O67825;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE GBM/WT RNA.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GBM;
RX MEDLINE=94076453; PubMed=8254770;
RA Graff J., Norman A., Feinstein S.M., Flehmig B.;
RT "Nucleotide sequence of wild-type hepatitis A virus GBM in comparison
to two cell culture adapted variants.";
RL J. Virol. 68:548-554(1994).
DR EMBL; X75215; CAAS3025.1; -
DR InterPro; IPR000408; Reg_chir_condens.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
PFam; PF00680; RNA_dep_RNA_pol; 1.
Pfam; PF00910; RNA_helicase; 1.
DR PROSITE; PS00626; RCCL_2; UNKNOWN 1.
FT CHAIN 1 791 P1 STRUCTURAL PROTEIN.
FT CHAIN 792 1422 P2 NONSTRUCTURAL PROTEIN.
FT CHAIN 1423 2227 P3 NONSTRUCTURAL PROTEIN.
SQ SEQUENCE 2227 AA; 251563 MW; 4C4D79D352F938B4 CRC64;

Query Match 99.4%; Score 1053; DB 12; Length 2227;
Best Local Similarity 98.5%; Pred. No. 3e-89; Indels 0; Gaps 0;
Matches 199; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECRKPYKEIRLEVGKORLKYAOEELSNEVLP PPRKKKGLFSQAK 60
DB 779 VDDPSEEDRRFRESHIECRKPYKEIRLEVGKORLKYAOEELSNEVLP PPRKKKGLFSQAK 838
QY 61 ISLFYTEEHEIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
DB 839 ISLFYTEEHEIMKFSWGRVGTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
QY 121 WTEMKDDKIVSLIEKFTSNKYSKYNPPHGMDLDEIANSKDFPMSETDLCFLLHMLN 180
DB 899 WTEMKDDKIVSLIEKFTSNKYSKYNPPHGMDLDEIANSKDFPMSETDLCFLLHMLN 958
QY 181 PKKINLADRMGLSGVQEIKEQ 202
DB 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 7
O67826 PRELIMINARY; PRT; 2227 AA.
AC O67826;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE GBM/HBS RNA.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GBM;

RX MEDLINE=94076453; PubMed=82524770;
 RA Grifff J., Normann A., Feinstein S.M., Flehmig B.;
 RT "Nucleotide sequence of wild-type hepatitis A virus GBM in comparison
 RT to two cell culture adapted variants.";
 RL J. Virol. 68:548-554 (1994).
 DR EMBL; X75216; CAA53026.1; -
 DR InterPro; IPR000408; Reg_chir_condens.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR Pfam; PF00680; RNA_dep_RNA_pol.1.
 DR Pfam; PF00910; RNA_helicase.1.
 DR PROSITE; PS00626; RCL1_2; UNKNOWN.1.
 FT CHAIN 1 791 P1 STRUCTURAL PROTEIN.
 FT CHAIN 792 1422 P2 NONSTRUCTURAL PROTEIN.
 FT CHAIN 1423 2227 P3 NONSTRUCTURAL PROTEIN.
 SQ SEQUENCE 2227 AA; 251496 MW; 488CB7C962319457 CRC64;

Query Match 99.4%; Score 1053; DB 12; Length 2227;
 Best Local Similarity 98.5%; Pred. No. 3e-89;
 Matches 199; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAOEELSNEVLPPPRKKKGLFSQAK 60
 DB 779 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAOEELSNEVLPPPRKKKGLFSQAK 838
 QY 61 ISLFYTEEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 839 ISLFYTEEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
 QY 121 WTEMKDDKIVSLIEKFTSNKYSKYNFPHGMDLDELIANSKDPFNMSETDLCFLHMLN 180
 DB 899 WTEMKDDKIVSLIEKFTSNKYSKYNFPHGMDLDELIANSKDPFNMSETDLCFLHMLN 958
 QY 181 PKKINLADRMGLSGVQEIKEQ 202
 DB 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 8

ID Q9WMA3 PRELIMINARY; PRT; 2227 AA.
 AC Q9WMA3;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Polyprotein.
 OS Hepatitis A virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OS Hepatovirus.
 NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AH2;
 RX MEDLINE=21386014; PubMed=11495028;
 RA Fujiwara K., Yokosuka O., Fukai K., Imaizaki F., Saitoh H., Omata M.;
 RT "Analysis of full-length hepatitis A virus genome in sera from
 RT patients with fulminant and self-limited acute type A hepatitis.";
 RL J. Hepatol. 35:112-119 (2001).
 DR EMBL; AB020565; BAA35103.1; -
 DR MEROPS; C03.005; -
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR000408; Reg_chir_condens.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR Pfam; PF00680; RNA_dep_RNA_pol.1.
 DR Pfam; PF00910; RNA_helicase.1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR PROSITE; PS00626; RCL1_2; UNKNOWN.1.
 SQ SEQUENCE 2227 AA; 251440 MW; E04F846CEC7F50PD CRC64;

Query Match 99.4%; Score 1053; DB 12; Length 2227;
 Best Local Similarity 98.5%; Pred. No. 3e-89;
 Matches 199; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAOEELSNEVLPPPRKKKGLFSQAK 60
 DB 779 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAOEELSNEVLPPPRKKKGLFSQAK 838
 QY 61 ISLFYTEEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 839 ISLFYTEEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
 QY 121 WTEMKDDKIVSLIEKFTSNKYSKYNFPHGMDLDELIANSKDPFNMSETDLCFLHMLN 180
 DB 899 WTEMKDDKIVSLIEKFTSNKYSKYNFPHGMDLDELIANSKDPFNMSETDLCFLHMLN 958
 QY 181 PKKINLADRMGLSGVQEIKEQ 202
 DB 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 9

ID Q91FH5 PRELIMINARY; PRT; 2227 AA.
 AC Q91FH5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Polyprotein.
 OS Hepatitis A virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OS Hepatovirus.
 NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAF-203;
 RA Baptista M.L., Silva M., de Lima M.A., Yoshida C.F., Gaspar A.M.,
 RA Pires Lopes M.Q., Galler R.;
 RT "Nucleotide sequence of the HAF-203 hepatitis A virus strain isolated
 RT in Brazil and expression of the VP1 gene in a bacterial system.";
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF268396; AAF80114.1; -
 DR MEROPS; C03.005; -
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR000408; Reg_chir_condens.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR Pfam; PF00680; RNA_dep_RNA_pol.1.
 DR Pfam; PF00910; RNA_helicase.1.
 DR PRINTS; PR00918; CALICVIRUSNS.
 DR PROSITE; PS00626; RCL1_2; UNKNOWN.1.
 SQ SEQUENCE 2227 AA; 251432 MW; 81913AEC8A04200 CRC64;

Query Match 99.4%; Score 1053; DB 12; Length 2227;
 Best Local Similarity 99.0%; Pred. No. 3e-89;
 Matches 200; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAOEELSNEVLPPPRKKKGLFSQAK 60
 DB 779 VDDPSEEDRRFRESHIECKRPYKELRLVGVKQRLKYAOEELSNEVLPPPRKKKGLFSQAK 838
 QY 61 ISLFYTEEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 839 ISLFYTEEHEIMKFSWGRVATDTRALRRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEK 898
 QY 121 WTEMKDDKIVSLIEKFTSNKYSKYNFPHGMDLDELIANSKDPFNMSETDLCFLHMLN 180
 DB 899 WTEMKDDKIVSLIEKFTSNKYSKYNFPHGMDLDELIANSKDPFNMSETDLCFLHMLN 958
 QY 181 PKKINLADRMGLSGVQEIKEQ 202
 DB 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 10

Q9WMA0

ID O9WMA0 PRELIMINARY; PRT: 2227 AA.
AC O9WMA0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyprotein.
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FH2;
RA MEDLINE=21386014; PubMed=11495028;
RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saito H., Omata M.;
RT "Analysis of full-length hepatitis A virus genome in sera from
RT patients with fulminant and self-limited acute type A hepatitis.";
RL J. Hepatol. 35:112-119(2001).
DR EMBL; AB020568; BAA35106.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000408; Reg_chir_condens.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_p3d.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR PROSITE; PS00626; RCCL_2; UNKNOWN 1.
SQ SEQUENCE 2227 AA; 25118 MW; 53B6B432127E9B CRC64;

Query Match 99.2%; Score 1050; DB 12; Length 2227;
Best Local Similarity 98.5%; Pred. No. 5.8e-89;
Matches 199; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRRPESHIECRKPYKELRLEVGKORLKYAOEELSNEVLPPPRKMGKGFPSQAK 60
Db 779 VDDPSEEDRRPESHIECRKPYKELRLEVGKORLKYAOEELSNEVSPPRKMGKGFPSQAK 838
Qy 61 ISLFYTEEHEIMKFSMRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRILRLNDEK 120
Db 839 ISLFYTEEHEIMKFSMRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRILRLNDEK 898
Qy 121 WTEMKDDKIIVSLIEKFTSNKYSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 180
Db 899 WTEMKDDKIIVSLIEKFTSNKYSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 958
Qy 181 PKKINLADRMGLSGVQEIKEQ 202
Db 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 11
O9WMA0 PRELIMINARY; PRT: 2227 AA.
AC O9WMA0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyprotein.
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FH3;
RA MEDLINE=21386014; PubMed=11495028;
RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saito H., Omata M.;
RT "Analysis of full-length hepatitis A virus genome in sera from
RT patients with fulminant and self-limited acute type A hepatitis.";
RL J. Hepatol. 35:112-119(2001).
DR EMBL; AB020569; BAA35107.1; -.
DR MEROPS; C03.005; -.

DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000408; Reg_chir_condens.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_p3d.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR PROSITE; PS00626; RCCL_2; UNKNOWN 1.
SQ SEQUENCE 2227 AA; 251387 MW; 817640187672A23D CRC64;

Query Match 98.9%; Score 1047; DB 12; Length 2227;
Best Local Similarity 98.5%; Pred. No. 1.1e-88;
Matches 199; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRRPESHIECRKPYKELRLEVGKORLKYAOEELSNEVLPPPRKMGKGFPSQAK 60
Db 779 VDDPSEEDRRPESHIECRKPYKELRLEVGKORLKYAOEELSNEVLPPPRKMGKGFPSQAK 838
Qy 61 ISLFYTEEHEIMKFSMRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRILRLNDEK 120
Db 839 ISLFYTEEHEIMKFSMRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRILRLNDEK 898
Qy 121 WTEMKDDKIIVSLIEKFTSNKYSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 180
Db 899 WTEMKDDKIIVSLIEKFTSNKYSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 958
Qy 181 PKKINLADRMGLSGVQEIKEQ 202
Db 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 12
O9WV03 PRELIMINARY; PRT: 2227 AA.
AC O9WV03;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyprotein.
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LJ38/WT;
RA Hu Y., Hu N.;
RT "Hepatitis A virus LJ38/WT";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF357222; AAM08224.1; -.
SQ SEQUENCE 2227 AA; 251379 MW; 975C48511E9213D7 CRC64;

Query Match 98.9%; Score 1047; DB 12; Length 2227;
Best Local Similarity 98.0%; Pred. No. 1.1e-88;
Matches 198; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPSEEDRRPESHIECRKPYKELRLEVGKORLKYAOEELSNEVLPPPRKMGKGFPSQAK 60
Db 779 VDDPSEEDRRPESHIECRKPYKELRLEVGKORLKYAOEELSNEVLPPPRKMGKGFPSQAK 838
Qy 61 ISLFYTEEHEIMKFSMRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRILRLNDEK 120
Db 839 ISLFYTEEHEIMKFSMRCVGTADTRALRRFGSLAAGRSVWTLMDAGVLTGRILRLNDEK 898
Qy 121 WTEMKDDKIIVSLIEKFTSNKYSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 180
Db 899 WTEMKDDKIIVSLIEKFTSNKYSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 958
Qy 181 PKKINLADRMGLSGVQEIKEQ 202
Db 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 13
 Q9WMA1 PRELIMINARY; PRT; 2227 AA.
 ID Q9WMA1
 AC Q9WMA1
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Polypeptide.
 OS Hepatitis A virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FH1
 RX MEDLINE=21386014; PubMed=11495028;
 RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saitoh H., Omata M.;
 RT "Analysis of full-length hepatitis A virus genome in sera from
 patients with fulminant and self-limited acute type A hepatitis.";
 RT J. Hepatol. 35:112-119(2001).
 EMBL: AB020567; BAA35105.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR000408; Reg_chir_condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00910; RNA_helicase.1.
 DR PRINTS: PR00918; CALICVIRUSNS.
 DR PROSITE: PS00626; RC01_2; UNKNOWN 1.
 DR PROSITE: PS00626; RC01_2; UNKNOWN 1.
 SQ SEQUENCE 2227 AA; 251415 MW; F92C8E2323FC5621 CRC64;

Query Match 98.8%; Score 1046; DB 12; Length 2227;
 Best Local Similarity 98.0%; Pred. No. 1.4e-88;
 Matches 199; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 VDDPSEEDRRFRESHIECKRPYKEIRLEVGKORLKYAQEELSNEVLP PPRKMKGLFSQAK 60
 DB 779 VDDPSEEDRRFRESHIECKRPYKEIRLEVGKORLKYAQEELSNEVLP PPRKMKGLFSQAK 838
 QY 61 ISLFYTEHEIMKFSWGVYADTRALRRFGSIAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 839 ISLFYTEHEIMKFSWGVYADTRALRRFGSIAAGRSVWTLMDAGVLTGRLIRLNDEK 898
 QY 121 WTEMKDDKIVSLIEKFTSNKYWSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 180
 DB 899 WTEMKDDKIVSLIEKFTSNKYWSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 958
 DB 181 PKKINLADRMGLSGVQEIKEQ 202
 DB 959 PKKINLADRMGLSGVQEIKEQ 980

RESULT 14
 Q67817 PRELIMINARY; PRT; 2218 AA.
 ID Q67817
 AC Q67817
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Complete genome.
 OS Hepatitis A virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F.G.
 RX MEDLINE=95381623; PubMed=7653108;
 RA Bendurec F., Pisani G., Divizia M., Pana A., Morace G.;
 RT "Complete nucleotide sequence of a cytopathic hepatitis A virus strain
 isolated in Italy.";
 RT Virus Res. 36:299-309(1995).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F.G.
 RA Morace G.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X83302; CA558281.1; -
 DR InterPro: IPR000408; Reg_chir_condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00910; RNA_helicase.1.
 DR PROSITE: PS00626; RC01_2; UNKNOWN 1.
 DR PROSITE: PS00626; RC01_2; UNKNOWN 1.
 SQ SEQUENCE 2218 AA; 250476 MW; 813B21D3E4E533CA CRC64;
 Query Match 98.6%; Score 1044; DB 12; Length 2218;
 Best Local Similarity 98.0%; Pred. No. 2.1e-88;
 Matches 198; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VDDPSEEDRRFRESHIECKRPYKEIRLEVGKORLKYAQEELSNEVLP PPRKMKGLFSQAK 60
 DB 773 VDDPSEEDRRFRESHIECKRPYKEIRLEVGKORLKYAQEELSNEVLP PPRKMKGLFSQAK 832
 QY 61 ISLFYTEHEIMKFSWGVYADTRALRRFGSIAAGRSVWTLMDAGVLTGRLIRLNDEK 120
 DB 833 ISLFYTEHEIMKFSWGVYADTRALRRFGSIAAGRSVWTLMDAGVLTGRLIRLNDEK 892
 QY 121 WTEMKDDKIVSLIEKFTSNKYWSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 180
 DB 893 WTEMKDDKIVSLIEKFTSNKYWSKVPFGMLDLEIANSKDPFNMSETDLCFLHMLN 952
 QY 181 PKKINLADRMGLSGVQEIKEQ 202
 DB 953 PKKINLADRMGLSGVQEIKEQ 974

RESULT 15
 Q9WMA4 PRELIMINARY; PRT; 2227 AA.
 ID Q9WMA4
 AC Q9WMA4
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Polypeptide.
 OS Hepatitis A virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AH1;
 RX MEDLINE=21386014; PubMed=11495028;
 RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saitoh H., Omata M.;
 RT "Analysis of full-length hepatitis A virus genome in sera from
 patients with fulminant and self-limited acute type A hepatitis.";
 RT J. Hepatol. 35:112-119(2001).
 EMBL: AB020564; BAA35102.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR000408; Reg_chir_condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00910; RNA_helicase.1.
 DR PRINTS: PR00918; CALICVIRUSNS.
 DR PROSITE: PS00626; RC01_2; UNKNOWN 1.
 DR PROSITE: PS00626; RC01_2; UNKNOWN 1.
 SQ SEQUENCE 2227 AA; 251304 MW; ODE6D2AEC29C0CE CRC64;
 Query Match 98.4%; Score 1042; DB 12; Length 2227;
 Best Local Similarity 98.0%; Pred. No. 3.3e-88;
 Matches 198; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VDDPSEEDRRFRESHIECKRPYKEIRLEVGKORLKYAQEELSNEVLP PPRKMKGLFSQAK 60

Db	779	VDDPRSEEDRRFESHIESRKPYPELPLEVGKORLKYAOEELSNEVLPPPRKIKGLFSQAK	838
Qy	61	ISLFYTEHEHEIMKFSMRGVADTRALRRPFSIAAGRSVWTLMDAGVLTGRILRLINDEX	120
Db	839	ISLFYTEHEHEIMKFSMRGVADTRALRRPFSMAAGRSVWTLMDAGVLTGRILRLINDEX	898
Qy	121	WTEMDDKIVSLIEKFTSNKYSKVNFPHGMLDLEIIAANSKDFPMSSETDLCFLHMLN	180
Db	899	WTEMDDKIVSLIEKFTSNKYSKVNFPHGMLDLEIIAANSKDFPMSSETDLCFLHMLN	958
Qy	181	PKKINLADRMGLSGVQEIKEQ	202
Db	959	PKKINLADRMGLSGVQEIKEQ	980

Search completed: April 2, 2003, 11:10:26
Job time : 61.2788 secs

THIS PAGE BLANK (USPTO)


```

FT Region 1520..1738
FT /label=3C
FT Region 1739..2227
FT /label=3D
XX US4894228-A.
XX 16-JAN-1990.
XX 12-JUL-1988; 88US-0217824.
XX 12-JUL-1988; 88US-0217824.
XX 12-JUL-1988; 88US-0652967.
XX (USSH ) US DEPT HEALTH & HUMAN.
XX Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstein SM,
PI Daemer RJ, Gust ID;
XX WPI: 1990-075557/10.
XX N-PSDB; AAO03512.
XX
XX Vaccine against hepatitis A virus infection - comprises novel
XX attenuated hepatitis A virus strain.
XX
XX Claim 1, Fig 1, 18pp; English.
XX
XX The attenuated HAV is useful for inducing protective immunity against
XX HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
XX several nucleotide changes distributed throughout the genome, is
XX attenuated for chimpanzees, elicits serum neutralising antibodies, and is
XX suitable for use as an HAV vaccine. It is noted that not all the changes
XX are necessary for attenuation and use as a vaccine.
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 990; DB 11; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.9e-93;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECKRKYKELRLEVQKRLKYAOEELSNEVLPPRKKMGKLFQAKISLFTYEEHEIMK 60
DB 792 SHIECKRKYKELRLEVQKRLKYAOEELSNEVLPPRKKMGKLFQAKISLFTYEEHEIMK 851
QY 61 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDKWTMDKIVSLI 120
DB 852 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDKWTMDKIVSLI 911
DB 121 EKFTSNKYKSNVFPFGMLDEEIAANSKDFPNMSETDLCFLHWINPKKINLADRMGL 180
DB 912 EKFTSNKYKSNVFPFGMLDEEIAANSKDFPNMSETDLCFLHWINPKKINLADRMGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980
XX
XX RESULT 2
XX AAW34074
XX ID AAW34074 standard; Protein; 2227 AA.
XX
XX AAW34074;
XX
XX 27-APR-1998 (first entry)
XX
XX Hepatitis A virus HM-175 protein sequence.
XX
XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
XX infection; vaccine.
XX
XX Hepatitis A virus HM-175.
XX
XX Key Location/Qualifiers

```

```

FT Protein 1..23
FT /label= VP4
FT 24..245
FT /label= VP2
FT 246..491
FT /label= VP3
FT 492..791
FT /label= VP1
FT 792..980
FT /label= 2A
FT 981..1087
FT /label= 2B
FT 1088..1422
FT /label= 2C
FT 1423..1496
FT /label= 3A
FT 1497..1519
FT /label= 3B
FT 1520..1738
FT /label= 3C
FT 1739..2227
FT /label= 3D
XX
XX WO9740166-A2.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06506.
XX
XX 19-APR-1996; 96US-0015642.
XX
XX (USSH ) US SEC DEPT HEALTH.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Emerson SU, Purcell RH, Raychaudhuri G;
XX
XX WPI: 1997-535850/49.
XX N-PSDB; AAT93023.
XX
XX Human attenuated HAV genome containing simian HAV 2C gene - useful
XX as vaccine against HAV infection
XX
XX Disclosure; Fig 13A-D; 66pp; English.
XX
XX This protein sequence is encoded by the human hepatitis A virus
XX (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
XX HAV/7 is obtained by passage of HM-175 in African Green Monkey
XX kidney cells. A claimed DNA construct (1) comprises a genome of
XX HAV, where the genome is a human attenuated HAV genome in which a
XX region of the 2C gene has been replaced by a corresponding region
XX from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
XX region of the 2C gene from AGM-27 contained in the construct
XX preferably encodes amino acids 120-328 of the 2C protein, amino
XX acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
XX transcript of (1); (2) a cell transfected with (1) or the RNA
XX transcript of (1); (3) a HAV genome as above; (4) antibodies to the
XX HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
XX its RNA transcript, can be used as a vaccine for preventing HAV in
XX a mammal. (1) or the RNA transcript can also be used to stimulate
XX the production of protective antibodies in the mammal.
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 990; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.9e-93;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECKRKYKELRLEVQKRLKYAOEELSNEVLPPRKKMGKLFQAKISLFTYEEHEIMK 60
DB 792 SHIECKRKYKELRLEVQKRLKYAOEELSNEVLPPRKKMGKLFQAKISLFTYEEHEIMK 851
QY 61 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDKWTMDKIVSLI 120

```

Db 852 FSWRGVTAADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMDKIVSLI 911
 Qy 121 EKFTSNKYWSKVPFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMIGL 180
 Db 912 EKFTSNKYWSKVPFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMIGL 971
 Qy 181 SGVQEIKEQ 189
 Db 972 SGVQEIKEQ 980

RESULT 3

AA18607
 ID AAB18607 standard; Protein; 2227 AA.

AA18607;
 AC AAB18607;
 XX
 XX 15-JAN-2001 (first entry)
 DT
 DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

HA; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

Hepatitis A virus.

US6113912-A.

05-SEP-2000.

07-JUN-1995; 95US-0475886.

18-SEP-1992; 92US-0947338.

17-SEP-1993; 93WO-US08610.

10-MAR-1995; 95US-0397232.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

WPI; 2000-586464/55.

N-PSDB; AAA75476.

Novel live hepatitis A virus adapted to growth in human fibroblast cell

infection, has modified genome compared to wild type

Disclosure; Fig 6A-K; 72pp; English.

The present sequence is derived from a wild type hepatitis A virus

(HAV) strain HM-174. The sequence is modified to produce HAV which

are adapted to growth in the human fibroblast-like cell line MRC-5.

The HAV is able to propagate in MRC-5 cells and retain appropriate

attenuation. It is useful as a live vaccine for prophylaxis of

hepatitis A in humans and other primates.

Sequence 2227 AA;

Query Match 100.0%; Score 990; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.9e-93;

Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKEIRLEVGQRLKYAOEELSNEVLPPRRKKKGIFSQAKISLFTTEHEIWK 60
 Db 792 SHIECRKPKYKEIRLEVGQRLKYAOEELSNEVLPPRRKKKGIFSQAKISLFTTEHEIWK 851
 Qy 61 FSWRGVTAADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMDKIVSLI 120
 Db 852 FSWRGVTAADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMDKIVSLI 911
 Qy 121 EKFTSNKYWSKVPFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMIGL 180
 Db 912 EKFTSNKYWSKVPFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMIGL 971

Qy 181 SGVQEIKEQ 189
 Db 972 SGVQEIKEQ 980

RESULT 4

AA18609
 ID AAB18609 standard; Protein; 2227 AA.

AA18609;
 AC AAB18609;
 XX
 XX 15-JAN-2001 (first entry)
 DT
 DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

HA; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

HAV 4380.

Hepatitis A virus.

US6113912-A.

05-SEP-2000.

07-JUN-1995; 95US-0475886.

18-SEP-1992; 92US-0947338.

17-SEP-1993; 93WO-US08610.

10-MAR-1995; 95US-0397232.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

WPI; 2000-586464/55.

N-PSDB; AAA75478.

Novel live hepatitis A virus adapted to growth in human fibroblast cell

infection, has modified genome compared to wild type

Disclosure; Columns 93-104; 72pp; English.

The present sequence is derived from a live attenuated hepatitis A

virus (HAV) of the invention, designated HAV 4380. The sequence is

produced by modifying wild type HAV strain HM-174. The HAV of the

invention are adapted to growth in the human fibroblast-like cell

line MRC-5. The HAV is able to propagate in MRC-5 cells and retain

appropriate attenuation. It is useful as a live vaccine for prophylaxis

of hepatitis A in humans and other primates.

Sequence 2227 AA;

Query Match 100.0%; Score 990; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.9e-93;

Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKEIRLEVGQRLKYAOEELSNEVLPPRRKKKGIFSQAKISLFTTEHEIWK 60
 Db 792 SHIECRKPKYKEIRLEVGQRLKYAOEELSNEVLPPRRKKKGIFSQAKISLFTTEHEIWK 851
 Qy 61 FSWRGVTAADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMDKIVSLI 120
 Db 852 FSWRGVTAADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMDKIVSLI 911
 Qy 121 EKFTSNKYWSKVPFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMIGL 180
 Db 912 EKFTSNKYWSKVPFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMIGL 971
 Qy 181 SGVQEIKEQ 189
 Db 972 SGVQEIKEQ 980

RESULT 5

AAE19899 standard; Protein; 2227 AA.

AAE19899;

18-JUN-2002 (first entry)

Hepatitis A virus (HAV) protein.

Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial; cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer.

Hepatitis A virus.

WO200213855-A2.

21-FEB-2002.

15-AUG-2001; 2001WO-1B01808.

17-AUG-2000; 2000US-225767P.

29-AUG-2000; 2000US-229175P.

03-NOV-2000; 2000US-0705547.

(TRIP-) TRIPEP AB.

Salberg M, Hultgren C;

WPI; 2002-241837/29.

N-PSDB; AAD31766.

Vaccine compositions for treating and preventing disease, preferably hepatitis C virus infection, comprises ribavirin and antigen that has epitope present in hepatitis C virus.

Claim 11; Page 82-87; 120pp; English.

The invention relates to a composition comprising ribavirin and an antigen preferably non structural 3 protein (NS3)/4A fragment of hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV sequence. The composition is useful for enhancing an immune response to a hepatitis C antigen in humans, domestic, sport or pet species and as vaccines for treating and preventing HCV infections. The composition is also useful for treating viral, bacterial, fungal diseases and cancer.

The present sequence is hepatitis A virus (HAV) protein.

Sequence 2227 AA;

Query Match 100.0%; Score 990; DB 23; Length 2227;

Best Local Similarity 100.0%; Pred. No. 4.9e-93;

Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SHIECKRKYKELRLEVGKQRLKYAOBELSNEVLPPPRKMGKLFQAKISLFYTEHEIMK 60

792 SHIECKRKYKELRLEVGKQRLKYAOBELSNEVLPPPRKMGKLFQAKISLFYTEHEIMK 851

61 FSWRGVYADTRALRRFGSLAAGRSVWTEMDAGVLTGRLIRLNDKWTMCKDKIVSLI 120

852 FSWRGVYADTRALRRFGSLAAGRSVWTEMDAGVLTGRLIRLNDKWTMCKDKIVSLI 911

121 EKFTSNKYKSNFPGMDLEELIANSKDFPNMSETDLCFLHMLNPKKINLADRMGL 180

912 EKFTSNKYKSNFPGMDLEELIANSKDFPNMSETDLCFLHMLNPKKINLADRMGL 971

181 SGVQEIKEQ 189

972 SGVQEIKEQ 980

RESULT 6

AAB18608

ID AAB18608 standard; Protein; 2227 AA.

AAB18608;

15-JAN-2001 (first entry)

Amino acid sequence of passage 35 Hepatitis A virus called P-35.

HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection; P-35 virus.

Hepatitis A virus.

US6113912-A.

05-SEP-2000.

07-JUN-1995; 95US-0475886.

18-SEP-1992; 92US-0947338.

17-SEP-1993; 93MO-US08610.

10-MAR-1995; 95US-0397232.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

WPI; 2000-586464/55.

N-PSDB; AAA75477.

Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.

Disclosure; Columns 67-78; 72pp; English.

The present sequence is derived from passage 35 of a wild type hepatitis A virus (HAV) strain HM-174. The resulting virus is designated P-35 virus. The sequence is modified to produce HAV which are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates.

Sequence 2227 AA;

Query Match 99.5%; Score 985; DB 21; Length 2227;

Best Local Similarity 99.5%; Pred. No. 1.6e-92;

Matches 188; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 SHIECKRKYKELRLEVGKQRLKYAOBELSNEVLPPPRKMGKLFQAKISLFYTEHEIMK 60

792 SHIECKRKYKELRLEVGKQRLKYAOBELSNEVLPPPRKMGKLFQAKISLFYTEHEIMK 851

61 FSWRGVYADTRALRRFGSLAAGRSVWTEMDAGVLTGRLIRLNDKWTMCKDKIVSLI 120

852 FSWRGVYADTRALRRFGSLAAGRSVWTEMDAGVLTGRLIRLNDKWTMCKDKIVSLI 911

121 EKFTSNKYKSNFPGMDLEELIANSKDFPNMSETDLCFLHMLNPKKINLADRMGL 180

912 EKFTSNKYKSNFPGMDLEELIANSKDFPNMSETDLCFLHMLNPKKINLADRMGL 971

181 SGVQEIKEQ 189

972 SGVQEIKEQ 980

RESULT 7

AAP60066 standard; Protein; 2227 AA.

AAP60066;

AAB18608

DT		26-JUN-1991 (first entry)
XX		
DE	Sequence of viral I _A 34 polypeptide encoded by the complete	
DE	nucleotide sequence of the HAV genome.	
XX		
XX	Diagnostics; vaccine; passive immunotherapy.	
OS	Hepatitis A virus.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..245
FT	Region	/label= P1.1A
FT	Region	246..491
FT	Region	/label= 1B
FT	Region	492..836
FT	Region	/label= 1C
FT	Region	837..980
FT	Region	/label= P2.2A
FT	Region	981..1076
FT	Region	/label= 2B
FT	Region	1077..1422
FT	Region	/label= 2C
FT	Region	1423..1484
FT	Region	/label= P3.3A
FT	Region	1485..1507
FT	Region	/label= 3B
FT	Region	1508..1678
FT	Region	/label= 3C
FT	Region	1679..2227
FT	Region	/label= 3D
XX		
XX	EPI99480-A.	
PN		
PD	29-OCT-1986.	
XX		
PF	03-APR-1986;	BEEP-0302465.
XX		
PR	03-APR-1985;	85US-0719329.
XX		
PA	(CHIR-) CHIRON CORP.	
XX		
P1	Dina D, Potter SJ, Vanneest GA, Caput D;	
XX		
DR	WPI; 1986-286213/44.	
N-P	PSDB; AAN60080.	
XX		
FT	Hepatitis A virus nucleotide sequence and polypeptide - and use	
FT	in prodn. of vaccines and diagnostic probes	
XX		
●	Claim 5; Fig 1; 18pp; English.	
CC	AAN60080 and oligonucleotide fragments are useful in detection of	
CC	hepatitis A virus; transformed hosts may be used for expression of	
CC	polypeptides and fragments useful in vaccines without risk of	
CC	infection by the virus or in prodn. of particles which are capable	
CC	of inducing immunocompetent B cells for passive immunotherapy. Pref.	
CC	epitope is derived from Aas 445-657 or 792-848 of the HAV	
CC	polypeptide sequence (AAN60066).	
XX		
SQ	Sequence	2227 AA;
Query Match	97.1%; Score 961; DB 7; Length 2227;	
Best Local Similarity	96.8%; Pred. No. 5e-90;	
Matches	183; Conservative 3; Mismatches 3; Indels 0; Gaps	0
OY	1 SHIECRKPYKEI.RLEVGKORLKYAOEELSNEVLPPEPKMKGLFSQAKISLFYTEHEIMK 60	
Db	792 THIESRKPYKEI.RLGKORLKYAOEELSNEVLPPEPKMKMGHFSQAKISLFYTEHEIMK 851	
OY	61 FSWRGVTADTRALRRGFGSLAAGRSVWTLEMAGVLTGLILNDEKMTMDDKTIVSLI 120	
Db	852 FSWRGVTADTRALRRGFGSLAAGRSVWTLEMAGVLTGLILNDEKMTMDDKTIVSLI 911	

```

Oy      121 EKFTSNKTSKYVPPHGLDLEETIAANSKDPPNMSSETDLCEFLHWLNPKINLADRMIGL 180
          |||
Db       912 EKFTSNKTSKYVPPHGLDLEETIAANSKDPPNMSSETDLCEFLHWLNPKINLADRMIGL 971

Oy      181 SGVOEIKEQ 189
          |||
Db       972 SGVOEIKEQ 980

RESULT 8
ID AAP50230 standard; Protein; 366 AA.
XX AAP50230;
AC AAP50230;
DT 28-NOV-1991 (first entry)
XX
DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).
XX
KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
XX diagnostic assay.
OS Hepatitis A virus.
XX
PN EPI38704-A.
PD 24-APR-1985.
PF 09-OCT-1984; 84EP-0402025.
PR 02-MAR-1984; 84US-0585942.
PR 14-OCT-1983; 83US-0541836.
PA (MERI ) MERCK & CO INC.
XX
PI Hughes JV, Scolnick EM, Tomassini JE;
XX
DR WPI; 1985-100818/17.
DR N-PSDB; AAN50274.
XX
PT New hepatitis A virus surface protein - useful for binding to
XX neutralising antibodies to the virus
XX
PS Claim 21; Page 46-48; 49pp; English.
CC VPI is isolated by solubilisation of the intact virus in an aq.
CC antonic surfactant and a reducing agent. The viral proteins are sepd.
CC and the protein of molecular wt. 33000 daltons is sepd.
XX
SQ Sequence 366 AA;

Query Match 34.6%; Score 343; DB 6; Length 366;
Best Local Similarity 98.5%; Pred. No. 4,8e-27;
Matches 65; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      1 SHIECRKPKEKLRLVEGKQRLKYAQBELISNEVLDPPEPMKGI,FSQAOKTSLFYTEBHEIMK 60
          |||
Db       301 SHIECRKPKEKLRLVEGKQRLKYAQBELISNEVLDPPEPMKGLFSQAOKTSLFYTEBHEIMK 360

Oy      61 FSWRGV 66
          |||
Db       361 FSWRGV 366

RESULT 9
ID AAP50116 standard; Protein; 993 AA.
XX AAP50116;
AC AAP50116;
DT 30-SEP-1991 (first entry)
XX
DE Sequence of Hepatitis A virus (HAV) immunogenic peptides
```

DE VP-1, VP-2, VP-3 and VP-4.
 XX Antigenic protein; immunogen; vaccine.
 XX Hepatitis A virus (strain CR326).
 OS EPI54587-A.
 PN 11-SEP-1985.
 XX 27-FEB-1985; 85EP-0400369.
 PF 02-MAR-1984; 84US-0585818.
 PR (MERI) MERCK & CO INC.
 PA Linemeyer DL, Menke JG, Reuben RG, Mitra SW,
 PI WPI; 1985-224964/37.
 XX N-PSDB; AAN50139.
 PT New nucleotide sequences coding for hepatitis A virus antigens -
 PT useful for eliciting normal immune response and in vaccines for
 PT protecting against the virus
 XX Example; Page 11-17; 32pp; English.
 XX Within the sequence in AAN50139 is encoded the information necessary
 CC to make the antigenic proteins of HAV. The sequences encoding for
 CC the structural proteins begin at base 403. The key sub-unit
 CC sequences within VP-1, designated Sequences I, II, III, IV, and V,
 CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
 CC nucleotide sequences which are valuable as encoding antigenic
 CC proteins are the sequences from base 1749 to base 2722. From base
 CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
 CC base 1749 to base 2722 is esp. valuable as a vector for producing
 CC antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the
 CC translation of a stop codon.
 SQ Sequence 993 AA;
 Query Match 34.6%; Score 343; DB 6; Length 993;
 Best Local Similarity 98.5%; Pred. No. 1.8e-26;
 Matches 65; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHIECKRPYKRLVGVKQRLKYAQBELSNVLPPIPRKMGKLFQAKISLFTTEHEIMK 60
 928 SHIECKRPYKRLVGVKQRLKYAQBELSNVLPPIPRKMGKLFQAKISLFTTEHEIMK 987
 QY 61 FSWRGV 66
 988 FSWRGV 993
 Db 988 FSWRGV 993
 RESULT 10
 AAP50231
 ID AAP50231 standard; Protein; 993 AA.
 XX AAP50231;
 AC 28-NOV-1991 (first entry)
 DT Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).
 XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KM diagnostic assay.
 XX Hepatitis A virus.
 OS Hepatitis A virus.
 PI Key
 FH Location/Qualifiers
 FT 628..993
 FT /note="claimed; X denotes translated stop codons

FT and unspecified triplets"
 XX EPI38704-A.
 XX 24-APR-1985.
 PD 09-OCT-1984; 84EP-0402025.
 XX 02-MAR-1984; 84US-0585942.
 PR 14-OCT-1983; 83US-0541836.
 XX (MERI) MERCK & CO INC.
 PA Hughes JV, Scolnick EM, Tomassini JE;
 PI WPI; 1985-100818/17.
 DR N-PSDB; AAN50274.
 XX New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus
 XX Disclosure; Page 17-23; 49pp; English.
 XX VP1 is isolated by solubilisation of the intact virus in an aq.
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.
 CC and the protein of molecular wt. 33000 daltons is sepd.
 SQ Sequence 993 AA;
 Query Match 34.6%; Score 343; DB 6; Length 993;
 Best Local Similarity 98.5%; Pred. No. 1.8e-26;
 Matches 65; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHIECKRPYKRLVGVKQRLKYAQBELSNVLPPIPRKMGKLFQAKISLFTTEHEIMK 60
 928 SHIECKRPYKRLVGVKQRLKYAQBELSNVLPPIPRKMGKLFQAKISLFTTEHEIMK 987
 QY 61 FSWRGV 66
 988 FSWRGV 993
 Db 988 FSWRGV 993
 RESULT 11
 AAP50287
 ID AAP50287 standard; Protein; 854 AA.
 XX AAP50287;
 AC 30-NOV-1991 (first entry)
 DT Sequence encoded by hepatitis A virus (HAV) cDNA from near the
 DE genome 5' terminus to the end of the area corresponding to the
 DE capsid protein region of poliovirus RNA.
 XX Hepatitis A virus assay; antigen; antibody.
 XX Hepatitis A virus.
 OS Hepatitis A virus.
 PN WO8501517-A.
 XX 11-APR-1985.
 PD 27-SEP-1984; 84MO-US01552.
 PF 30-SEP-1983; 83US-0537911.
 XX (MASI) MASSACHUSETTS INST TECH.
 PA Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 PI Racanelli VR;
 XX WPI; 1985-098846/16.
 DR N-PSDB; AAN50330.


```

XX New hepatitis A virus CDNA - useful in assays for the virus and
PT for prodn. of the viral antigen and antibodies to it
XX
XX Example; Fig 7; 60pp; English.
XX
CC The inventors claim HAV cDNA and a method for producing it, whereby
CC large amts. can be obtd. economically. The cDNA is useful in the
CC assay for detection of HAV quickly and easily and with high
CC sensitivity and specificity. The HAV cDNA is also used in the prodn.
CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.
XX
SQ Sequence 854 AA;

Query Match 33.5%; Score 332; DB 6; Length 854;
Best Local Similarity 100.0%; Pred. No. 2e-25;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKEIRLEVGKORLKYAOEISNEVLPPEPRKMGKLFQAKISLFTYEEHEIMK 60
792 SHIECRKPKYKEIRLEVGKORLKYAOEISNEVLPPEPRKMGKLFQAKISLFTYEEHEIMK 851
OY 61 FSW 63
DB 852 FSW 854

RESULT 12
AAW95559
ID AAW95559 standard; Protein; 1077 AA.
XX
XX AAW95559;
XX
DT 28-APR-1999 (first entry)
XX
XX A partial hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus protein; HAV; P2 region;
XX
XX cell-culture-adapted HAV strain; infection; accelerated growth.
XX
XX Hepatitis A virus.
XX
XX US5849562-A.
XX
XX 15-DEC-1998.
XX
XX 06-JUN-1995; 95US-0468926.
XX
XX 06-NOV-1991; 91US-0788262.
XX
XX 30-SEP-1983; 83US-0537911.
XX
XX 27-SEP-1984; 84US-0654942.
XX
XX 06-OCT-1988; 88US-0256135.
XX
XX 06-JUN-1995; 95US-0468926.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Emerson SU, Purcell RH;
XX
XX WPI; 1999-094412/08.
XX
XX N-PSDB; AAX01006.
XX
XX Chimeric hepatitis A virus strains - with P2 region from
PT cell-culture-adapted strain in wild-type genome
XX
XX
XX Disclosure; Fig 7A-L; 36pp; English.
XX
XX The present sequence represents a partial hepatitis A virus (HAV)
XX protein. The specification describes a DNA construct consisting
XX of a wild-type HAV genome in which the P2 region is replaced by the
XX P2 region from a cell-culture-adapted HAV strain. The construct is
XX used to demonstrate that mutations in the P2 region of a
XX cell-culture-adapted HAV strain are sufficient for establishment of
XX infection and accelerated growth in cell culture.

```

```

XX SQ Sequence 1077 AA;
XX
XX Query Match 33.5%; Score 332; DB 20; Length 1077;
XX Best Local Similarity 100.0%; Pred. No. 2.8e-25;
XX Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKEIRLEVGKORLKYAOEISNEVLPPEPRKMGKLFQAKISLFTYEEHEIMK 60
DB 1015 SHIECRKPKYKEIRLEVGKORLKYAOEISNEVLPPEPRKMGKLFQAKISLFTYEEHEIMK 1074
OY 61 FSW 63
DB 1075 FSW 1077

RESULT 13
AAR32426
ID AAR32426 standard; Protein; 1091 AA.
XX
XX AAR32426;
XX
DT 17-DEC-2001 (updated)
DT 10-JUN-1993 (first entry)
XX
XX Translated from 5' region of Hepatitis A Virus genomic clone.
XX
XX HAV HM-175; chronic liver disease; picornavirus.
XX
XX Hepatitis A virus.
XX
XX Key Location/Qualifiers
XX 238..1091
XX Region /label=ORF
XX /note="second putative initiation codon at
XX position 240"
XX Region 1..711
XX /note="X's correspond to nonsense codons,
XX i.e. this region is not an ORF"
XX
XX USN7788262-N.
XX
XX 15-DEC-1992.
XX
XX 30-SEP-1983; 83US-0536911.
XX
XX 27-SEP-1984; 84US-0654942.
XX
XX 06-OCT-1988; 88US-0256135.
XX
XX 30-SEP-1983; 83US-0536911.
XX
XX 06-NOV-1991; 91US-0788262.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Baltimore D, Feinstone SM;
XX Purcell RH, Racanelli VR, Ticehurst JR;
XX
XX WPI; 1993-067429/08.
XX
XX N-PSDB; AAO36934.
XX
XX Hepatitis A virus CDNA prodn. - for diagnostic use and for prodn.
PT of antigen and antibodies
XX
XX Disclosure; Fig 7; 65pp; English.
XX
XX HAV virion RNA was extracted from the livers of marmosets which had
XX been inoculated with HAV (the HAV had previously been passaged twice
XX in marmosets). The RNA was used to prepare ds cDNA clones by
XX standard methods. Clones contg. inserts which hybridised to RNA from
XX HAV-infected African Green Monkey Kidney cells were selected for
XX further analysis. A 7.4kb restriction map (about 9% of the HAV
XX genome) was constructed from 5 overlapping inserts. The sequence of
XX the first 3.3kb (approx.) from the 5'-terminus was determined. An
XX amino acid sequence was deduced from the entire clone and an open

```

CC reading frame was identified starting at position 238. A comparison
 CC of the predicted HAV amino acid sequences with the known capsid
 CC protein sequences of other picornaviruses (poliovirus, foot and
 CC mouth disease virus and encephalomyelitis virus) revealed areas of
 CC local homology.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government-owned NTIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)

CC Sequence 1091 AA;

Query Match 33.5%; Score 332; DB 14; Length 1091;

Best Local Similarity 100.0%; Pred. No. 2.8e-25;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRPPYKELRLVGVKQRLKYAQBELSNELVPPPRKMKGLFSQAKISLFTYEEHEIMK 60

1029 SHIECKRPPYKELRLVGVKQRLKYAQBELSNELVPPPRKMKGLFSQAKISLFTYEEHEIMK 1088

61 FSW 63

DB 1089 FSW 1091

RESULT 14

AAR15629

AC AAR15629;

DT 17-DEC-2001 (updated)

DT 17-MAR-1992 (first entry)

DE Capsid region of cyno-HAV isolate CY-145.

XX Hepatitis A virus; cynomolgus; HAV; monkey; vaccine; macaque.

OS Cynomolgus monkey hepatitis A virus, isolate CY-145.

XX Key Location/Qualifiers

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

XX Disclosure; Fig 3; 23pp; English.

PS The sequence was deduced from the nucleotide sequence obt. by PCR
 CC amplification of cyno-HAV viral RNA obt. from the stool of a
 CC cynomolgus monkey with serologically and histologically confirmed
 CC spontaneous hepatitis A. The sequence differs from the human HAV
 CC isolate HM175 (Cohen, J.I., et al. (1987) Proc. Natl. Acad. Sci.
 CC USA 84, 2497-2501), mainly in the VP3 and VP1 proteins. The Gln-Val
 CC pair at the VP3-VP1 cleavage site in the human isolate is replaced
 CC by a Gln-Thr pair in the cyno-HAV. The other two cleavage sites are
 CC the same. Two residues have been identified as part of the immuno-
 CC dominant region (see feature table) and are different to those in
 CC the same position in human HAV. The protein and peptides derived
 CC from it can be used in the prepn. of vaccines for the prevention of
 CC HAV infection.

CC See also AAR15056.

CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government-owned NTIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)

CC Sequence 839 AA;

Query Match 21.7%; Score 214.5; DB 12; Length 839;

Best Local Similarity 85.7%; Pred. No. 2.9e-13;

Matches 42; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

QY 1 SHIECKRPPYKELRLVGVKQRLKYAQBELSNELVPPPRKMKGLFSQAKIS 49

DB 792 SHIECKRPPYKELRLVGVKQRLKYAQBELSNELVPPPRKMKGLFSQAKIS 839

RESULT 15

AAM42930

AC AAM42930;

DT 28-APR-1998 (first entry)

DE Immunogenic Hepatitis A virus peptide YK-1665.

XX Immunogenic peptide; immunogenic epitope; P2A protein;

XX immune response; antibody.

OS Synthetic.

OS Hepatitis A virus.

XX WO9740147-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US06891.

XX 19-APR-1996; 96US-0015644.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX FIELDS HA, Knudskov YE;

XX WPI; 1997-535831/49.

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an

XX immune response to HAV in a mammal or to detect the presence of

XX antibodies against HAV in a mammal

XX Claim 18; Page 112; 140pp; English.

XX Peptides AAM42922-30 are immunogenic peptides corresponding to

XX immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are

XX substantially similar to a portion of the amino acid sequence of the P2A

XX protein of HAV corresponding to amino acids 792-980. Compositions

CC containing the peptides can be used to induce an immune response to HAV
CC in a mammal. The peptides can also be used to detect the presence of
CC antibodies against HAV in mammalian serum. The peptides can also be used
CC to make an antibody against HAV by administering the peptide to a
CC mammal.

XX	Sequence	25 AA;
SQ		

Sequence 25 AA;

Query Match	13.0%	Score 129;	DB 18;	Length 25;
Best Local Similarity	100.0%	Pred. No. 2e-06;		
Matches 25; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 19 QRLKYAQEELSNEVLPPPRKMG^{LF} 43
|||
Db 1 QRLKYAQEELSNEVLPPPRKMG^{LF} 25

Search completed: April 2, 2003, 11:07:57
Job time : 55.688 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 2, 2003, 11:05:59 ; Search time 11.1176 Seconds
(without alignments)
500.190 Million cell updates/sec

Title: US-10-104-966-12_COPY_792_980
Perfect score: 990
Sequence: 1 SHIECRKPKYKELRLVGVGQKOR.....KINLADRMGLSGVQEIKEQ 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	990	100.0	2227	3	US-08-475-886-2
2	990	100.0	2227	3	US-08-475-886-6
3	990	100.0	2227	4	US-08-397-232-2
4	990	100.0	2227	4	US-08-397-232-4
5	990	100.0	2227	4	US-09-171-387-2
6	990	100.0	2227	4	US-09-653-499-2
7	990	100.0	2227	4	US-09-653-499-6
8	985	99.5	2227	3	US-08-475-886-4
9	985	99.5	2227	4	US-09-653-499-4
10	332	33.5	1091	6	551630-2
11	214.5	21.7	839	1	US-08-087-016-2
12	84.5	8.5	1353	4	US-09-398-193-99
13	82.5	8.3	643	2	US-08-245-511-47
14	82.5	8.3	1294	4	US-08-600-993A-47
15	82.5	8.3	1294	4	US-09-473-717-2
16	81	8.2	229	4	US-09-248-528-17
17	81	8.2	229	4	US-09-549-108-17
18	81	8.2	229	4	US-09-549-111-17
19	81	8.2	229	4	US-09-549-116-17
20	81	8.2	229	4	US-09-550-394-17
21	79.5	8.0	1305	1	US-08-441-139-9
22	79.5	8.0	1305	4	US-08-864-785-3
23	79.5	8.0	1353	3	US-08-894-173-2
24	79.5	8.0	1353	4	US-09-398-193-2
25	79.5	8.0	1353	4	US-09-473-717-3
26	76	7.7	582	4	US-08-194-560-2
27	76	7.7	3443	2	US-08-416-603-2

28	74	7.5	400	4	US-08-961-083-190	Sequence 190, App
29	73.5	7.4	513	4	US-09-724-519-6	Sequence 6, Appli
30	73.5	7.4	513	4	US-09-592-037-6	Sequence 8, Appli
31	73.5	7.4	575	4	US-09-724-519-8	Sequence 8, Appli
32	73.5	7.4	575	4	US-09-592-037-8	Sequence 8, Appli
33	73	7.4	600	4	US-09-423-468A-13	Sequence 13, Appli
34	73	7.4	1440	4	US-09-357-251-37	Sequence 37, Appli
35	73	7.4	1512	4	US-09-443-184-48	Sequence 48, Appli
36	72	7.3	1259	4	US-09-134-001C-3757	Sequence 3757, Ap
37	70.5	7.1	158	4	US-09-134-001C-4358	Sequence 4358, Ap
38	70.5	7.1	897	4	US-09-134-001C-3600	Sequence 3600, Ap
39	70.5	7.1	976	4	US-09-104-324B-4	Sequence 4, Appli
40	70.5	7.1	1053	4	US-09-724-519-2	Sequence 2, Appli
41	70.5	7.1	1053	4	US-09-592-037-2	Sequence 2, Appli
42	70.5	7.1	1447	4	US-09-376-330-17	Sequence 17, Appli
43	70	7.1	1057	4	US-08-928-862-3	Sequence 3, Appli
44	69.5	7.0	1057	4	US-09-541-782-10	Sequence 10, Appli
45	69.5	7.0	1057	4	US-09-723-820-10	Sequence 10, Appli

ALIGNMENTS

```
RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475, 886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947, 338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397, 232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match 100.0%; Score 990; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLVGVGQKORLKYAOEELSNBYLPPPRKMGKGFSGAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKELRLVGVGQKORLKYAOEELSNBYLPPPRKMGKGFSGAKISLFTTEHEIMK 851
QY 61 FSWRGVTVADTRALRRFGFSLAAGRSVWTLTENDAGVLTGRLIRLNDKTEMDKDIYSLI 120
DB 852 FSWRGVTVADTRALRRFGFSLAAGRSVWTLTENDAGVLTGRLIRLNDKTEMDKDIYSLI 911
QY 121 EKFTSNKTKWSKNFPHGMJDEEIIAANSKDPFNMSSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKTKWSKNFPHGMJDEEIIAANSKDPFNMSSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVQEIKEQ 189
DB 972 SGVQEIKEQ 980
RESULT 2
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
```

```
/ GENERAL INFORMATION:
/ APPLICANT: FUNKHOUSER, ANN W
/ APPLICANT: EMERSON, SUZANNE U
/ APPLICANT: PURCELL, ROBERT H
/ APPLICANT: D'HONDT, ERIC
/ TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
/ FILE REFERENCE: 20264262US2
/ CURRENT APPLICATION NUMBER: US/08/475,886A
/ CURRENT FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 07/947,338
/ EARLIER FILING DATE: 1992-09-18
/ EARLIER APPLICATION NUMBER: 08/397,232
/ EARLIER FILING DATE: 1995-03-10
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 2227
/ TYPE: PRT
/ ORGANISM: Attenuated (4380) HAV, strain HM-175
/ DB-475-886-6

Query Match          100.0%; Score 990; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
QY 121 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 3
US-08-397-232-2
/ Sequence 2, Application US/08397232A
/ Patent No. 6180110
/ GENERAL INFORMATION:
/ APPLICANT: FUNKHOUSER, ANN W
/ APPLICANT: EMERSON, SUZANNE U
/ APPLICANT: PURCELL, ROBERT H
/ APPLICANT: D'HONDT, ERIC
/ TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
/ FILE REFERENCE: 20264262US1
/ CURRENT APPLICATION NUMBER: US/08/397,232A
/ CURRENT FILING DATE: 1995-04-17
/ EARLIER APPLICATION NUMBER: 07/947,338
/ EARLIER FILING DATE: 1992-09-18
/ EARLIER APPLICATION NUMBER: PCT/US93/08610
/ EARLIER FILING DATE: 1993-09-17
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 2227
/ TYPE: PRT
/ ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
/ US-08-397-232-2

Query Match          100.0%; Score 990; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
QY 121 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 4
US-08-397-232-4
/ Sequence 4, Application US/08397232A
/ Patent No. 6180110
/ GENERAL INFORMATION:
/ APPLICANT: FUNKHOUSER, ANN W
/ APPLICANT: EMERSON, SUZANNE U
/ APPLICANT: PURCELL, ROBERT H
/ APPLICANT: D'HONDT, ERIC
/ TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
/ FILE REFERENCE: 20264262US1
/ CURRENT APPLICATION NUMBER: US/08/397,232A
/ CURRENT FILING DATE: 1995-04-17
/ EARLIER APPLICATION NUMBER: 07/947,338
/ EARLIER FILING DATE: 1992-09-18
/ EARLIER APPLICATION NUMBER: PCT/US93/08610
/ EARLIER FILING DATE: 1993-09-17
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 2227
/ TYPE: PRT
/ ORGANISM: Attenuated (4380) HAV, strain HM-175
/ US-08-397-232-4

Query Match          100.0%; Score 990; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
QY 121 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 5
US-09-171-387-2
/ Sequence 2, Application US/09171387
/ Patent No. 6280734
/ GENERAL INFORMATION:
/ APPLICANT: RAYCHAUDHURI, GOPA;
/ APPLICANT: EMERSON, SUZANNE, U.;
/ APPLICANT: PURCELL, ROBERT, H.;
/ TITLE OF INVENTION: SIMIAN-HUMAN HAV
/ HAVING A CHIMERIC 2C PROTEIN
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESSES:
```

```
DB 792 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
QY 121 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 4
US-08-397-232-4
/ Sequence 4, Application US/08397232A
/ Patent No. 6180110
/ GENERAL INFORMATION:
/ APPLICANT: FUNKHOUSER, ANN W
/ APPLICANT: EMERSON, SUZANNE U
/ APPLICANT: PURCELL, ROBERT H
/ APPLICANT: D'HONDT, ERIC
/ TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
/ FILE REFERENCE: 20264262US1
/ CURRENT APPLICATION NUMBER: US/08/397,232A
/ CURRENT FILING DATE: 1995-04-17
/ EARLIER APPLICATION NUMBER: 07/947,338
/ EARLIER FILING DATE: 1992-09-18
/ EARLIER APPLICATION NUMBER: PCT/US93/08610
/ EARLIER FILING DATE: 1993-09-17
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 2227
/ TYPE: PRT
/ ORGANISM: Attenuated (4380) HAV, strain HM-175
/ US-08-397-232-4

Query Match          100.0%; Score 990; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECKRKYKELRLVGVKORLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
QY 121 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYKSKNFPFGMDLDEIANSKDPFNMSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 5
US-09-171-387-2
/ Sequence 2, Application US/09171387
/ Patent No. 6280734
/ GENERAL INFORMATION:
/ APPLICANT: RAYCHAUDHURI, GOPA;
/ APPLICANT: EMERSON, SUZANNE, U.;
/ APPLICANT: PURCELL, ROBERT, H.;
/ TITLE OF INVENTION: SIMIAN-HUMAN HAV
/ HAVING A CHIMERIC 2C PROTEIN
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESSES:
```

ADDRESSEE: MORGAN & PINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996

ATTORNEY/AGENT INFORMATION:
NAME: William S. Feiler
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229US1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-171-387-2

Query Match 100.0%; Score 990; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKEIRLEVGQRKLYKAOEELSNVLPPEPRKMGKGFSGAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKEIRLEVGQRKLYKAOEELSNVLPPEPRKMGKGFSGAKISLFTTEHEIMK 851

QY 61 FSWRGVTDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 911

QY 121 EKFTSNKYWSKYNPFHGMLDLEIANSKDPNNSETDLCFLHLWLNPKKINLADRMIGL 180
DB 912 EKFTSNKYWSKYNPFHGMLDLEIANSKDPNNSETDLCFLHLWLNPKKINLADRMIGL 971

QY 181 SGVGEIKEQ 189
DB 972 SGVGEIKEQ 980

RESULT 6
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18

PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 990; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKEIRLEVGQRKLYKAOEELSNVLPPEPRKMGKGFSGAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKEIRLEVGQRKLYKAOEELSNVLPPEPRKMGKGFSGAKISLFTTEHEIMK 851

QY 61 FSWRGVTDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 911

QY 121 EKFTSNKYWSKYNPFHGMLDLEIANSKDPNNSETDLCFLHLWLNPKKINLADRMIGL 180
DB 912 EKFTSNKYWSKYNPFHGMLDLEIANSKDPNNSETDLCFLHLWLNPKKINLADRMIGL 971

QY 181 SGVGEIKEQ 189
DB 972 SGVGEIKEQ 980

RESULT 7
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 990; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-107;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKEIRLEVGQRKLYKAOEELSNVLPPEPRKMGKGFSGAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKEIRLEVGQRKLYKAOEELSNVLPPEPRKMGKGFSGAKISLFTTEHEIMK 851

QY 61 FSWRGVTDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 911

QY 121 EKFTSNKYWSKYNPFHGMLDLEIANSKDPNNSETDLCFLHLWLNPKKINLADRMIGL 180
DB 912 EKFTSNKYWSKYNPFHGMLDLEIANSKDPNNSETDLCFLHLWLNPKKINLADRMIGL 971

QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 8

US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 99.5%; Score 985; DB 3; Length 2227;
Best Local Similarity 99.5%; Pred. No. 4,7e-107;
Matches 188; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRPYKELRLLEVQKQRLKYAOEELSNEVLPPPRKMGKLFQSAKISLFYTEEHEIMK 60
DB 792 SHIECKRPYKELRLLEVQKQRLKYAOEELSNEVLPPPRKMGKLFQSAKISLFYTEEHEIMK 851
QY 61 FSWRGVATDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVATDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
QY 121 EKFTSNKYWSKYNFPHGMLDLEIJAANSKDPNNSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYWSKYNFPHGMLDLEIJAANSKDPNNSETDLCFLHMLNPKKINLADRMGL 971
DB 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 9

US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 99.5%; Score 985; DB 4; Length 2227;
Best Local Similarity 99.5%; Pred. No. 4,7e-107;
Matches 188; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRPYKELRLLEVQKQRLKYAOEELSNEVLPPPRKMGKLFQSAKISLFYTEEHEIMK 60
DB 792 SHIECKRPYKELRLLEVQKQRLKYAOEELSNEVLPPPRKMGKLFQSAKISLFYTEEHEIMK 851
QY 61 FSWRGVATDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVATDTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
QY 121 EKFTSNKYWSKYNFPHGMLDLEIJAANSKDPNNSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYWSKYNFPHGMLDLEIJAANSKDPNNSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 10

5516630-2
Patent No. 5516630
APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
BAROUDY, BAHIGE M.
TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/788,262
FILING DATE: 06-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 256,135
FILING DATE: 06-OCT-1988
APPLICATION NUMBER: 654,942
FILING DATE: 27-SEP-1984
APPLICATION NUMBER: 537,911
FILING DATE: 30-SEP-1983
SEQ ID NO: 2
LENGTH: 1091
5516630-2

Query Match 33.5%; Score 332; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 2,8e-30;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKRPYKELRLLEVQKQRLKYAOEELSNEVLPPPRKMGKLFQSAKISLFYTEEHEIMK 60
DB 1029 SHIECKRPYKELRLLEVQKQRLKYAOEELSNEVLPPPRKMGKLFQSAKISLFYTEEHEIMK 1088
QY 61 FSW 63
DB 1089 FSW 1091

RESULT 11

US-08-087-016-2
Sequence 2, Application US/08087016
Patent No. 5430135
GENERAL INFORMATION:
APPLICANT: NAINAN, OMANA V.
APPLICANT: MARGOLIS, HAROLD S.
APPLICANT: ROBERTSON, BETTY H.
APPLICANT: BRINTON, MARGO H.
APPLICANT: EBBERT, JAMES W.
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1615 L Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,016
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,828
FILING DATE: 03-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 839 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-087-016-2

Query Match 21.7%; Score 214.5; DB 1; Length 839;
Best Local Similarity 85.7%; Pred. No. 1.5e-16;
Matches 42; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

QY 1 SHIERKPYKELRLEVKGORLKYAOEELSNEVLPPPRKKGLFSQAKIS 49
DB 792 SHIE-KKPYKELRLEVKGORFYAREELSNEILPPPRKLGKLFSSQKIS 839

RESULT 12
US-09-398-193-99
Sequence 99, Application US/09398193
Patent No. 6197581
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: Adenylate cyclase and uses therefor
FILE REFERENCE: P24360-
CURRENT APPLICATION NUMBER: US/09/398,193
CURRENT FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 104
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 99
LENGTH: 1353
TYPE: PRT
ORGANISM: Human
US-09-398-193-99

Query Match 8.5%; Score 84.5; DB 4; Length 1353;
Best Local Similarity 23.5%; Pred. No. 0.71;
Matches 54; Conservative 30; Mismatches 83; Indels 63; Gaps 12;

QY 3 IECRPYKELRLE-----VGKORLKYAOEELSNEV-----LPPPRKKGLFSQAKISL 50
DB 323 LEVFALEKRMHWSVPRILADLMDKGDSESENSKRIATSPKRRKKSSIQKAPLAF 382
QY 51 -FYTEEHIMKFSMRGVTADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDERK 108

DB 383 RPFKMOQIEVSI L FADIVGFTK-----MSANKSAHALVGLDLFGRPDLCE--- 431
QY 109 TEMKDKIVSLIEKFTSNKYSKVNFPH-----GMLD-LEBIANSKDPFN 153
DB 432 -ETKCKEISTL-----GDCYVACGCEPRADHAYCCEIEMGIMKAIQFQCKEMWN 485
QY 154 M-----SETDLCFLH-----WLNPKKINLADM--LGLSGVOEIKE 188
DB 486 MRGVHTRTVLCGLGMRRFRKFDVWSN--DVNLAMLMQLGVAQKVIHSE 533

RESULT 13
US-08-245-511-47
Sequence 47, Application US/08245511
Patent No. 5928900
GENERAL INFORMATION:
APPLICANT: Masure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-245-511-47

Query Match 8.3%; Score 82.5; DB 2; Length 643;
Best Local Similarity 26.1%; Pred. No. 0.42;
Matches 42; Conservative 19; Mismatches 53; Indels 47; Gaps 9;

QY 21 LKYAOEELSNEVLPPPRKKGLFSQAKISLFTYEEHIMKFSMRGVTADTRALRRFGPS 79
DB 107 LKYADKXSDALYVOESIKGLDAYK-----GEIKDSQVGIKALDEQTVG---YT 155
QY 80 LAAGRSVWTLMDAGVLTGRLIRLNDERKTEWKD-----KIVSLIEK- 122
DB 156 INKPSPFNNSKTTMGVLA-----PVNEEFINSKGDPAKATDPSSILYNGPYLLKSVTKS 211
QY 123 ---FTSN-KYMSKV-----PPHGMLEDEFLAANSKD 150
DB 212 SVEFAKNPNYMDKDVHIDKYGLSFWDQ-DTSKPAENFKD 251

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 2, 2003, 11:10:39 ; Search time 21.2685 Seconds

(without alignments)
543.277 Million cell updates/sec

Title: US-10-104-966-12_COPY_792_980

Perfect score: 990
Sequence: 1 SHIECRKPKYKELRLVGRQR.....KINLADRLMLGLSGVEIKERQ 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PTCT_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	2227	9 US-10-104-966-12	Sequence 12, Appl
2	990	100.0	2227	9 US-10-135-988-2	Sequence 2, Appl
3	990	100.0	2227	9 US-10-135-988-6	Sequence 6, Appl
4	990	100.0	2227	10 US-09-929-955-12	Sequence 12, Appl
5	985	99.5	2227	9 US-10-135-988-4	Sequence 4, Appl
6	84.5	8.5	1353	10 US-09-751-100B-99	Sequence 99, Appl
7	82.5	8.3	1294	12 US-10-071-223-2	Sequence 2, Appl
8	79.5	8.0	501	9 US-10-066-521-2	Sequence 364, App
9	79.5	8.0	1250	10 US-09-801-368-364	Sequence 2, Appl
10	79.5	8.0	1353	10 US-09-751-100B-2	Sequence 3, Appl
11	79.5	8.0	1353	12 US-10-071-223-3	Sequence 3, Appl
12	78	7.9	1204	10 US-09-935-291A-9	Sequence 8, Appl
13	77	7.8	454	10 US-09-866-582-8	Sequence 10292, A
14	75.5	7.6	297	10 US-09-815-242-10292	Sequence 13742, A
15	75.5	7.6	297	10 US-09-815-242-13742	Sequence 13636, A
16	75	7.5	810	10 US-09-815-242-13636	Sequence 190, App
17	74	7.5	400	10 US-09-765-272-190	Sequence 4, Appl
18	74	7.5	419	9 US-10-182-263-4	Sequence 5, Appl
19	73	7.4	419	9 US-10-182-263-5	

20	73	7.4	419	9 US-10-182-263-6	Sequence 6, Appl
21	71	7.2	419	9 US-10-182-263-3	Sequence 3, Appl
22	69.5	7.0	300	9 US-10-278-173-34	Sequence 34, Appl
23	69.5	7.0	538	10 US-09-758-269-8	Sequence 8, Appl
24	69.5	7.0	538	10 US-09-758-269-18	Sequence 18, Appl
25	69.5	7.0	538	10 US-09-758-269-33	Sequence 33, Appl
26	69	7.0	478	9 US-09-866-050A-512	Sequence 512, App
27	68	6.9	316	10 US-09-815-242-11179	Sequence 11179, A
28	68	6.9	362	10 US-09-740-288A-34	Sequence 34, Appl
29	68	6.9	363	10 US-09-740-288A-35	Sequence 35, Appl
30	68	6.9	628	10 US-09-881-752A-144	Sequence 144, App
31	68	6.9	870	12 US-10-078-929-188	Sequence 188, App
32	67.5	6.8	450	9 US-09-738-626-4815	Sequence 4815, App
33	67	6.8	243	10 US-09-765-272-220	Sequence 220, App
34	67	6.8	243	10 US-09-901-106-2	Sequence 2, Appl
35	66.5	6.7	241	9 US-09-738-626-6844	Sequence 6844, App
36	66.5	6.7	396	10 US-09-801-368-138	Sequence 138, App
37	66	6.7	171	10 US-09-764-853-452	Sequence 452, App
38	66	6.7	253	10 US-09-815-242-11740	Sequence 514, App
39	66	6.7	346	12 US-10-007-693-70	Sequence 70, Appl
40	66	6.7	413	9 US-09-738-626-4157	Sequence 4157, App
41	66	6.7	419	9 US-10-182-263-1	Sequence 1, Appl
42	66	6.7	419	9 US-09-978-917A-4	Sequence 4, Appl
43	66	6.7	461	9 US-10-182-263-2	Sequence 2, Appl
44	66	6.7	461	9 US-09-978-917A-2	
45	66	6.7	461	9 US-09-978-917A-2	

ALIGNMENTS

RESULT 1				
US-10-104-966-12				
Sequence 12, Application US/10104966				
Patent No. US2002015124A1				
GENERAL INFORMATION:				
APPLICANT: Matci Saliberg				
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND				
TITLE OF INVENTION: METHODS OF USE THEREOF				
FILE REFERENCE: TRIPEP-23AUSCI				
CURRENT APPLICATION NUMBER: US/10/104.966				
CURRENT FILING DATE: 2002-03-22				
PRIOR APPLICATION NUMBER: 09/705,547				
PRIOR FILING DATE: 2000-11-03				
PRIOR APPLICATION NUMBER: 60/229,175				
NUMBER OF SEQ ID NOS: 15				
SOFTWARE: FastSeq for Windows Version 4.0				
SEQ ID NO 12				
LENGTH: 2227				
TYPE: PRT				
ORGANISM: Artificial Sequence				
FEATURE:				
OTHER INFORMATION: Hepatitis A virus sequence				
US-10-104-966-12				
Query Match				
Best Local Similarity 100.0%; Pred. No. 7.7e-98;				
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Length 2227;				
QY	1	SHIECRKPKYKELRLVGRQR	100.0%; DB 9;	Length 2227;
DB	792	SHIECRKPKYKELRLVGRQR	100.0%; DB 9;	Length 2227;
QY	61	PSMRGVADTRALRRFGSLAAGRSVWTLMDAGVLTGLRLINDEKMTMDKIVSLI	100.0%; DB 9;	Length 2227;
DB	852	PSMRGVADTRALRRFGSLAAGRSVWTLMDAGVLTGLRLINDEKMTMDKIVSLI	100.0%; DB 9;	Length 2227;
QY	121	EKFTSNKTKWSKNFPHGMLDEEIANSKDPFNMSETDLCFLHNLNPKKINLADRLML	100.0%; DB 9;	Length 2227;
DB	912	EKFTSNKTKWSKNFPHGMLDEEIANSKDPFNMSETDLCFLHNLNPKKINLADRLML	100.0%; DB 9;	Length 2227;

QY	181	SGVQEIKEQ	189
Db	972	SGVQEIKEQ	980

RESULT 2
US-10-135-988-2

```

; Sequence 2, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:

```

; APPLICANT: PUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

Query Match	100.0%	Score 990	DB 9	Length 2227
Best Local Similarity	100.0%	Pred. NO. 7.7e-98		
Matches 189; Conservative	0	Mismatches 0	Indels 0	Gaps 0

Qy 61 FSWRGVTTADTRALRRFGFSLAAGHSVWTTTMDAGVLTGRLRLNDEKWTMKDKIYSLI 1200

Db 852 FSWRGVTTADTRALRRFGFSLAAGHSVWTTTMDAGVLTGRLRLNDEKWTMKDKIYSLI 911

Oy 121 EKFTSNKYWSKYNFPHGMLDLEELAAANSKDPPNMSSETDLCLLHWLNPKKLINADRMGL 180

Db 912 EKFTSNKYWSKYNFPHGMLDLEELAAANSKDPPNMSSETDLCLLHWLNPKKLINADRMGL 971

181 SGVQEIKEQ 189
972 SGVQEIKEQ 980

RESULT 3
US-10-13

US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1

; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; APPLICANT: D'HONDT, ERIC

ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match	100.0%;	Score 990;	DB 9;	length 2227;
Best Local Similarity	100.0%;	Pred. No. 7.7e-98;		
Matches 189;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy 61 FSWMGVTAIDRALRRRGFSLAAGRSVWTLMDAGVLTGLIRLNDEKWTENKDDKIVSLI 120

Db 852 FSWMGVTAIDRALRRRGFSLAAGRSVWTLMDAGVLTGLIRLNDEKWTENKDDKIVSLI 911

```
Oy      121 EKFSNKKYWSKVNPFGMLDDEEIAANSKDFPNMSETDLCFLHLHLPKKKNLADRMGL 180C
          |||||
Db       912 EKFSNNKYWSKVPFPGMLDDEEIAANSKDFPNMSETDLCFLHLHLPKKKNLADRMGL 977T
```

QY	181	SGVQEIKEQ	189
Db	972	SGVQEIKEQ	980

RESULT 4
US-09-929-955-12

US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1

```

1  GENERAL INFORMATION:
2  APPLICANT:       Matci Sallberg
3  APPLICANT:       Catharina Hultgren
4  TITLE OF INVENTION:  VACCINES CONTAINING RIBAVIRIN AND
5  TITLE OF INVENTION:  METHODS OF USE THEREOF
6  INVENTOR:         Matci Sallberg and Catharina Hultgren

```

OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match	100.0%;	Score 990;	DB 10;	Length 2227;
Best Local Similarity	100.0%;	Pred. No. 7.7e-98;		
Matches 189;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy 61 FSWRGVYADPTALRRFFGSLAAGRSVWLTLEMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 852 FSWRGVYADPTALRRFFGSLAAGRSVWLTLEMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911

Oy 121 EKFTSKYKWSKYNEPFGHMLDEEIAANSKDFPNMSETDLCFLHWLNPKKINLADRMGL 188
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 912 EKFTSKYKWSKYNEPFGHMLDEEIAANSKDFPNMSETDLCFLHWLNPKKINLADRMGL 977

QY	181	SGVQEIKEQ	185
Db	972	SGVQEIKEQ	980

RESULT 5
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 99.5%; Score 985; DB 9; Length 2227;
Best Local Similarity 99.5%; Pred. No. 2.7e-97;
Matches 188; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKLELVGKQRLKYAOEELSNEVLPFRKKKGLFSQAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKLELVGKQRLKYAOEELSNEVLPFRKKKGLFSQAKISLFTTEHEIMK 851
QY 61 FSWRGVATDTRALRRFGFSLAAGRSVWTEMDAGVLTGRLIRLNDKMTKEMDDKIVSLI 120
DB 852 FSWRGVATDTRALRRFGFSLAAGRSVWTEMDAGVLTGRLIRLNDKMTKEMDDKIVSLI 911
QY 121 EKFTSNKYWSKVNFPHGMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMGL 180
DB 912 EKFTSNKYWSKVNFPHGMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMGL 971
QY 181 SGVQRIKEX 189
DB 972 SGVQRIKEX 980

RESULT 6
US-09-751-100B-99
; Sequence 99, Application US/09751100B
; Patent No. US20020142436A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor
; FILE REFERENCE: P27948A
; CURRENT APPLICATION NUMBER: US/09/751,100B
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 99
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-100B-99

Query Match 8.5%; Score 84.5; DB 10; Length 1353;
Best Local Similarity 23.5%; Pred. No. 2.8;
Matches 54; Conservative 30; Mismatches 83; Indels 63; Gaps 12;

QY 3 IECRKPKYKLELVGKQRLKYAOEELSNEVLPFRKKKGLFSQAKISL 50
DB 323 LEVAKLKERMHVMPRIADDLKQDEESENSVKRIATSSPKRRKKSSIQKAPIAF 382
QY 51 --FYTEHEIMKFSWRGVATDTRALRRFGFSLAAGRSVWTEMDAGVLTGRLIRLNDK 108

DB 383 RPFKMQQIEEVSILFADIVGFTRK-----MSANKSAHALVGLNDLFGDRDRICE--- 431
QY 109 TEMKDKIVSLIEKFTSNKYWSKVNFPH-----GMLD-LEEIAANSKDFPN 153
DB 432 -ETKCEKISTL-----GDCYCVAGCPBRADHAYCCIEEMGLMKAIEQFOCKEKEMVN 485
QY 154 M-----SETDLCFLH-----WLNPKKINLADRM--LGLSGVQEIKE 188
DB 486 MRVGHVHTGTVLGGLGMRRFKFDVMSN--DVLNLMLEQLGVAQKVHISE 533

RESULT 7
US-10-071-223-2
; Sequence 2, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
; APPLICANT: Storm, Daniel R.
; APPLICANT: Hacker, Beth
; APPLICANT: Tomlinson, James E.
; APPLICANT: COR Therapeutics, Inc.
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
; FILE REFERENCE: 44481-5029-02-US
; CURRENT APPLICATION NUMBER: US/10/071,223
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: 09/473,717
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PCT/US98/13541
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/098,559
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 08/886,440
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1294
; TYPE: PRT
; ORGANISM: human type IX adenylyl cyclase
US-10-071-223-2

Query Match 8.3%; Score 82.5; DB 12; Length 1294;
Best Local Similarity 23.5%; Pred. No. 4.3;
Matches 54; Conservative 30; Mismatches 83; Indels 63; Gaps 12;

QY 3 IECRKPKYKLELVGKQRLKYAOEELSNEVLPFRKKKGLFSQAKISL 50
DB 323 LEVAKLKERMHVMPRIADDLKQDEESENSVKRIATSSPKRRKKSSIQKAPIAF 382
QY 51 --FYTEHEIMKFSWRGVATDTRALRRFGFSLAAGRSVWTEMDAGVLTGRLIRLNDK 108
DB 383 RPFKMQQIEEVSILFADIVGFTRK-----MSANKSAHALVGLNDLFGDRDRICE--- 431
QY 109 TEMKDKIVSLIEKFTSNKYWSKVNFPH-----GMLD-LEEIAANSKDFPN 153
DB 432 -ETKCEKISTL-----GDCYCVAGCPBRADHAYCCIEEMGLMKAIEQFOCKEKEMVN 485
QY 154 M-----SETDLCFLH-----WLNPKKINLADRM--LGLSGVQEIKE 188
DB 486 MRVGHVHTGTVLGGLGMRRFKFDVMSN--DVLNLMLEQLGVAQKVHISE 533

RESULT 8
US-10-066-521-2
; Sequence 2, Application US/10066521
; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR

```

1  TITLE OF INVENTION:  PROTEIN FAMILY AND USES THEREOF
2  FILE REFERENCE:  07334-334001
3
4  CURRENT APPLICATION NUMBER:  US/10/066,521
5
6  CURRENT FILING DATE:  2002-06-25
7
8  PRIOR APPLICATION NUMBER:  60/318,645
9
10 PRIOR FILING DATE:  2001-09-10
11
12 PRIOR APPLICATION NUMBER:  60/265,231
13
14 PRIOR FILING DATE:  2001-01-31
15
16 NUMBER OF SEQ. ID NOS:  25
17
18 SOFTWARE:  FastSeq for Windows Version 4.0
19
20 SEQ. ID NO. 2
21
22 LENGTH:  501
23
24 TYPE:  PRT
25
26 ORGANISM:  Homo sapiens
27
28 US-10-066-521-2

```

Query Match	8.0%	Score 79.5	DB 9	Length 501
Best Local Similarity	25.8%	Pred. No. 2.5		
Matches 41, Conservative	27	Mismatches 58	Indels 33	Gaps 9

```

Dc      12 LREVGKQRL-KYAOEELSNELPPRKMGLFS--QAQSLPFTBEE---EIKKFPMWG 65
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      75 LFLQINRDQLMTKAOEENRNKLNTPRKRMKETPOLWEKETCLHVPBFYEYTKMNEYKE 134
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Cy      66 VTADTRALRRFGSFAAGRSVWTLMDAGVLIGRLIR-----LNDEKTEMKDKIV- 117
        ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      135 LND-----AYTAAARRHTVLEGGPGIGKTTILRKVMLDMAEGNLMKDSINEKLTVW 186
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Cy      118 -SLIEKTSNKYWSKVNFPHGMLDLEEPAANSKDPPNNS 155
        :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      187 RELCMGFITNK----NQI-IIDMEN--TSLDDPSIA 215

```

```

RESULT 9
US-09-801-368-364
: Sequence 364, Application US/09801368
: Patent No. US20020128250A1
: GENERAL INFORMATION:
: APPLICANT: Busby, Robert
: APPLICANT: Call, Brian
: APPLICANT: Hecht, Peter
: APPLICANT: Holtzman, Doug
: APPLICANT: Madden, Kevin
: APPLICANT: Maxon, Mary
: APPLICANT: Milne, Todd
: APPLICANT: No. US20020128250A1man, Thea
: APPLICANT: Royer, John
: APPLICANT: Salama, Sofie
: APPLICANT: Sherman, Amir
: APPLICANT: Silva, Jeff
: APPLICANT: Sumners, Eric
: TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
: FILE REFERENCE: 109272.147
: CURRENT APPLICATION NUMBER: US/09/801,368
: PRIOR APPLICATION NUMBER: US 09/487,558
: PRIOR FILING DATE: 2000-01-19
: PRIOR APPLICATION NUMBER: US 60/160,587
: PRIOR FILING DATE: 1999-10-20
: NUMBER OF SEQ ID NOS: 440
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 364
: LENGTH: 1250
: TYPE: prt
: ORGANISM: Saccharomyces cerevisiae
US-09-801-368-364

```

```

Query Match      8.0%; Score 79.5; DB 10; Length 1250;
Best Local Similarity 21.0%; Pred. No. 8.6;
Matches 52; Conservative 28; Mismatches 65; Indels 103; Gaps 12.

QY  28 LSNVEUPP--PRCKGLFSQAKI---SLFY-----TEHEHINKFS---W----- 63
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

```

Dh 659 LSNBYLDQKPNQKXKPSFQPLPLTAESLXYRRNNTDTEYNNIFAISLQWSEFALHVRN 718
Oy 64 -----RGVADPRALRRFG-----FSLAGR 84
Dh 719 NGNGTLELGHVVDVTSIHIEGSSVDRARRSSAVFMPOKLVNLLPQSFNDELSTLAPGK 778
Oy 85 S-----VWTLLEM-----DAGVLGTGLRLRN-DEKWTMCKDKIVSLIEKFTS 125
Dh 779 ESATLSVYVTLTDSSTLRKSTKSTVGEESTISPENISLQBDEKJSTGSPTSYLSVTOEJAR 836
Oy 126 NKYYSKXVFPBGMGLDLEBIANSKDFPMMSTBDLCFLHMLNPKFI---NLADRMGLS 188
Dh 839 SFYARRINDPATL-----LPTLS-----LLESLDDEKXVNDILNDRITGEV 881
Oy 182 GVQEIKEQ 189
Dh 882 VINEIKR 889

```

```

RESULT 10
US-09-751-100B-2
; Sequence 2, Application US/09751100B
; Patent No. US20020142436A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: Human Adenylate Cyclase and Use Therefor
; FILE REFERENCE: P27948A
; CURRENT APPLICATION NUMBER: US/09/751,100B
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1353
; TYPE: prt
; ORGANISM: Mouse
US-09-751-100B-2

```

Query Match	8.0%;	Score 79.5;	DB 10;	Length 1353;
Best Local Similarity	23.0%;	Pred. NO. 9.6;		
Matches	53;	Conservative	31;	Mismatches 83;
				Indels 63;
				Gaps 12;

```

QY 3 IECKRPEYELALE-----VCKQKLTXYAQEGLSNEY-----LPPPKMKGLFSQAISL 50
Db 323 LEVEFALTEERHIMSVPRRIIADDLIMKOCDESEBSVVRHATSSPKYKKKSSIOKATIAF 3022
QY 51 --FYTEHEIEMKFSWRCGATADTRALRRFGESLAAGRSVWLTLEMDAGVLTSLIRLNDKMW 108
Db 383 RFPKMQQOIEEBSIILPADI VGF TK-----MSANKSAHLVGLINDLPFRFDR LCE--- 433
QY 109 TEMKDKIIVSLIEKFTSNKYKSVKNFPH-----GMULD-LEELIANSKQFPN 15
Db 432 -QTKCEKISTL-----GDYCYVAGCEPRRADHAYCCIEMLGMIKAI EOPCECKEMVN 485
QY 154 M-----SETDLCFLH-----WLNKPKINLADRM--LGLSGVGEI KE 188
Db 486 MEVGVHVTGVLGGLIGMRREKFDVWSN--DVNLNLNLEQGLVACKVHISE 533

```

```

RESULT 11
US-10-071-223-3
; Sequence 3, Application US/10071223
; Patent No. US20020137174A1
; GENERAL INFORMATION:
; APPLICANT: Storm, Daniel R.
; APPLICANT: Hacker, Beth
; APPLICANT: Tomlinson, James E.
; APPLICANT: COR Therapeutics, Inc.
; APPLICANT: University of Washington
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF A HUMAN ADENYLYL
; TITLE OF INVENTION: CYCLASE
; FILE REFERENCE: 44481-5029-02-US
; CURRENT APPLICATION NUMBER: US/10/071,223
; CURRENT FILING DATE: 2002-02-11

```

```

; PRIOR APPLICATION NUMBER: 09/473,717
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PCT/US98/13541
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/098,559
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 08/886,440
; PRIOR FILING DATE: 1997-07-01
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: murine type IX adenylyl cyclase
US-10-071-223-3
```

```

Query Match      8.0%; Score 79.5; DB 12; Length 1353;
Best Local Similarity 23.0%; Pred. No. 9.6;
Matches 53; Conservative 31; Mismatches 83; Indels 63; Gaps 12;
```

```

3 IECRKPYKELRL-----VGKQRLKYAOEELSNEY-----LPPRKMGKGFSAQKISL 50
323 LEVEKALKERMLHSWPRRIADLMKQDEESENSVKRHATSPKRRKKSSIQKPIAF 382
51 --FYTBHEIMKFSWRCVTAADTRALRRFGFSLAAGRSVTLMDAGVLTGRLIRLNDKRW 108
383 RPFKMQOIEEVSILFADIVGFTK-----MSANKSAHALVGLNDLFGFEDRLCE--- 431
109 TEMKDKIVSLIEKFTSNKYMSKANPPIH-----GMLD-LEETIANSKDFPN 153
432 -QTKCEKISTL-----GDCYYCVAGCPPEPRADHAYCCIEMLGIMTIRAIQFOCEKKEMVN 485
154 M-----SETDLCFLH-----WLPKKINLADRM--LGSVGOEIKE 188
486 MGVGHTGTGVLGCGILGMRKFKEDWSN--DVNLINMEQLGVAAGKHISE 533
```

```

RESULT 12
US-09-935-291A-9
; Sequence 9, Application US/09935291A
; Patent No. US20020107373A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: 49937, 49931, AND 49933, NOVEL HUMAN TRANSPORTER
; TITLE OF INVENTION: FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MNT-184
; CURRENT APPLICATION NUMBER: US/09/935,291A
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,504
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: 60/250,932
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 9
; LENGTH: 1204
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-291A-9
```

```

Query Match      7.9%; Score 78; DB 10; Length 1204;
Best Local Similarity 21.4%; Pred. No. 12;
Matches 36; Conservative 29; Mismatches 61; Indels 42; Gaps 7;
```

```

32 VLPERRKMK-----GLFSQAKISLFYTE-----EHEIMKFSWRC-VTAADTRALRR 75
488 VVPELPIETLSLAVNTSLALAKLWYCTEPPRIPFAGKVEVCCFDKGTSLTSDSLVVG 547
76 RGFSLAAGRSVW-----TLEMDAGVLTGRLIR--LNDKFTMKMD 114
548 VAGLDGKEVTPVSSIPVETRHAIASCHSLMQLDGLTGVDPLEKAMLTAVDWTLTKDE 606
```

```

Qy 115 KIVSLIEKFTSNKYMSKVNPPHGMLEETIANSKDFPMSETDLCFL 162
Db 607 KVFPSRIKIQGLKIHQRFHFASALKRMVSLASYEK-----LGSTDLICYI 650
```

```

RESULT 13
US-09-866-582-8
; Sequence 8, Application US/09866582
; Patent No. US20020127620A1
; GENERAL INFORMATION:
; APPLICANT: Witman, George B.
; APPLICANT: Pazour, Gregory J.
; APPLICANT: Rosenbaum, Joel L.
; APPLICANT: Cole, Douglas G.
; TITLE OF INVENTION: INTRAFACELLAR TRANSPORT
; FILE REFERENCE: 07917-145001
; CURRENT APPLICATION NUMBER: US/09/866,582
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/206,923
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Chlamydomonas reinhardtii
US-09-866-582-8
```

```

Query Match      7.8%; Score 77; DB 10; Length 454;
Best Local Similarity 23.3%; Pred. No. 4;
Matches 28; Conservative 22; Mismatches 38; Indels 32; Gaps 6;
```

```

Qy 80 LAAGSVWTLMDAGVLT-----TGRILRLN-----DEKWTMKD-KIVSLIEKFTSNKY 128
Db 193 LSSGIAIYPMNRPVGAHVAHQPCYGRIVAGSCAMPDDKMLDKENSKINDPFFKLE--- 249
129 WSKVNPFGMLDLEETIANSKDFPMSETDLCFLHMLNPKKINLADRM-LGSVGOEIKE 188
250 -----PHSKIQGLNDIDAEPPVSDLK-----LLPDRASLADKLKGC--LQEIID 291
```

```

RESULT 14
US-09-815-242-10292
; Sequence 10292, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
```

```
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10292
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10292
```

```
Query Match          7.6%; Score 75.5; DB 10; Length 297;
Best Local Similarity 24.3%; Pred. No. 3.3;
Matches 43; Conservative 24; Mismatches 69; Indels 41; Gaps 7;
```

```
QY 19 ORLKVAQEELSENL---PPPRKKGLEFSQAKISLFYTEHEHEIMKFSWRG--VTADTRAL 73
DB 37 ORIKOLENMFQGPLVTRVTPRPTE---QGGKLLALLRQVELLEBEMLGDEQTSSTPLL 92
QY 74 RRFGLSLAAGR---SVWTLMDAGVLTGRLIRLN---DEKWTMK-----112
DB 93 ----LSLAVNADSLATWLLPALAPVLADSPIRLNLQVEDETRTQERLRGRGEVVGAVSIQH 148
113 -----DDKIYSLIEKFTSNKYMSKVNFPFGMLDLEETIANSKDFPMNSETDLCFL 162
DB 149 QALPSCLVDKLGALDYLFFVASKPFAERYFPNGVTRSSILKAPAVAFDHLDDHQAFLL 205
```

RESULT 15

```
US-09-815-242-13742
; Sequence 13742, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cair, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.01A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13742
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13742
```

```
Query Match          7.6%; Score 75.5; DB 10; Length 297;
Best Local Similarity 23.7%; Pred. No. 3.3;
Matches 42; Conservative 26; Mismatches 68; Indels 41; Gaps 7;
```

```
QY 19 ORLKVAQEELSENL---PPPRKKGLEFSQAKISLFYTEHEHEIMKFSWRG--VTADTRAL 73
DB 37 ORIKOLENMFQGPLVTRVTPRPTE---QGGKLLALLRQVELLEBEMLGDEQTSSTPLL 92
```

```
QY 74 RRFGLSLAAGR---SVWTLMDAGVLTGRLIRLN---DEKWTMK-----112
DB 93 ----LSLAVNADSLATWLLPALAPVLADSPIRLNLQVEDETRTQERLRGRGEVVGAVSIQH 148
QY 113 -----DDKIYSLIEKFTSNKYMSKVNFPFGMLDLEETIANSKDFPMNSETDLCFL 162
DB 149 QALPSCLVDKLGALDYLFFVASKPFAERYFPNGVTRSSILKAPAVAFDHLDDHQAFLL 205
```

```
Search completed: April 2, 2003, 11:22:20
Job time : 25.2685 secs
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 11:05:19 ; Search time 25.6189 Seconds
(without alignment)
709.218 Million cell updates/sec

Title: US-10-104-966-12_COPY_792_980
Perfect score: 990
Sequence: 1 SHIECKRKYKLELVGSKQR.....KINLADRMGLSGVQEIKEQ 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
1 number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	2227	1 GNNYHM	genome polypotein
2	990	100.0	2227	1 GNNYHM	genome polypotein
3	985	99.5	2227	1 GNNYMK	genome polypotein
4	984	99.4	2227	1 GNNYHB	genome polypotein
5	911	92.0	2230	1 GNNYSA	genome polypotein
6	333	33.6	1358	2 A03905	genome polypotein
7	317	32.0	852	1 GNNYHA	genome polypotein
8	236	23.8	341	2 S04137	genome polypotein
9	214.5	21.7	839	1 GNNYS2	genome polypotein
10	144	14.5	56	2 P00427	genome polypotein
11	143	14.4	56	2 P00429	genome polypotein
12	139	14.0	56	2 P00430	genome polypotein
13	137	13.8	56	2 P00431	genome polypotein
14	136	13.7	56	2 P00432	genome polypotein
15	135	13.6	56	2 P00433	genome polypotein
16	134	13.5	56	2 P00434	genome polypotein
17	126	12.7	55	2 P00435	genome polypotein
18	120	12.1	56	2 P00436	genome polypotein
19	119	12.0	56	2 P00437	genome polypotein
20	115.5	11.7	55	2 P00438	genome polypotein
21	102	10.3	319	2 JH0135	genome polypotein
22	93	9.4	330	2 A03267	hypothetical prote
23	90.5	9.1	309	2 H71089	hypothetical prote
24	86.5	8.7	588	2 E87521	peptidase M1 fam1
25	86	8.7	1289	1 RMXR3	mRNA guanylyltrans
26	85	8.6	381	2 T18876	hypothetical prote
27	83.5	8.4	610	2 T06590	galactonolactone d
28	82.5	8.3	643	2 S54395	permease-like prot
29	82	8.3	387	2 G87537	acyl-CoA dehydroge

30	81.5	8.2	278	2 B82388	probable maltose o
31	81	8.2	529	2 G90286	hypothetical prote
32	80.5	8.1	1034	2 S50917	aminomethyltransfe
33	80	8.1	570	2 A45249	alpha-glucosidase
34	79.5	8.0	355	2 A70185	conserved hypotet
35	79.5	8.0	1006	2 S20126	exoribonuclease RA
36	79.5	8.0	1250	2 A39578	SSD1 protein - yea
37	79.5	8.0	1353	2 UC4279	adenylate cyclase
38	79.5	8.0	1374	2 A71724	dna-directed RNA p
39	79	8.0	555	2 H96762	hypothetical prote
40	79	8.0	1019	2 T50251	hypothetical colle
41	79	8.0	1363	2 T15653	hypothetical prote
42	78.5	7.9	235	2 G90224	hypothetical prote
43	78.5	7.9	374	2 AD1970	serine/threonine k
44	78.5	7.9	1010	2 B86259	protein T12C24.2 f
45	78	7.9	276	2 E75036	3-methyladenine DN

ALIGNMENTS

RESULT 1

GNNYHM
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A25981
R:COhen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with di
A:Reference number: A25981; MUID:87061253; PMID:3023706
A:Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M4707; NID:G329582; PID:AAA45465.1; PID:G329583
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-245/Product: coat protein 1B #status predicted <VP2>
F:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP1>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1518/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cytosine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 990, DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-79;
Matches 189, Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SHIECKRKYKLELVGSKQR.....KINLADRMGLSGVQEIKEQ 189	DB	792	SHIECKRKYKLELVGSKQR.....KINLADRMGLSGVQEIKEQ 189
Qy	61	FSWGVATDTRALRRFGFSLAAGRSVWTLNDAGVLTGRLRLNDEKTEKMDKIVSLI 120	DB	852	FSWGVATDTRALRRFGFSLAAGRSVWTLNDAGVLTGRLRLNDEKTEKMDKIVSLI 120
Qy	121	EKFTSNKTKWSKVNPPHGMULDIEITAANSKDPNNSETDLCFLHMLNPKKINLADRMGL 180	DB	912	EKFTSNKTKWSKVNPPHGMULDIEITAANSKDPNNSETDLCFLHMLNPKKINLADRMGL 180
Qy	181	SGVQEIKEQ 189	DB	972	SGVQEIKEQ 980

Qy 121 EKFTSNKYKWSKVNPFHGMJLDEEIAANSKDPFNNSETDLCFLHMLNPKKINLADRMJGL 180
Db 912 EKFTSNKYKWSKVNPFHGMJLDEEIAANSKDPFNNSETDLCFLHMLNPKKINLADRMJGL 971
Qy 181 SGVOEIKEQ 189
Db 972 SGVOEIKEQ 980

RESULT 5

GENMYA

genome polyprotein - simian hepatitis A virus (strain AGM-27)

N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro

C/Species: simian hepatitis A virus

C/Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 16-Jun-2000

C/Accession: A30470; S04885; S03965

R/Tsarev, S.A.

Submitted to JIPID, April 1991

A/Reference number: A30470

A/Accession: A30470

A/Molecule type: genomic RNA

A/Residues: 1-2230 <TSA>

A/Cross-references: GB:D00924; NID:G222597; PIDN:BA00766.1; PID:G222598

R/Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.

J. Gen. Virol. 72, 1677-1683, 1991

A/Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure at

A/Reference number: J01080; MUID:91311420; PMID:1649901

A/Content: annotation

A/Note: neither amino acid nor nucleotide sequence is given

R/Balayan, M.S.; Kusov, Y.Y.; Andjapardize, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik

submitted to the EMBL Data Library, May 1989

A/Reference number: S04885

A/Accession: S04885

A/Molecule type: genomic RNA

A/Residues: 1750-2164 <BAL1>

A/Cross-references: EMBL:X15461; NID:961971; PIDN:CAA33490.1; PID:9330268

R/Balayan, M.S.; Kusov, Y.Y.; Andjapardize, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik

PEBS Lett. 247, 425-428, 1989

A/Title: Variations in genome fragments coding for RNA polymerase in human and simian he

A/Reference number: S03965; MUID:85232168; PMID:2541023

A/Accession: S03965

A/Molecule type: genomic RNA

A/Residues: 1960-2164 <BAL2>

A/Cross-references: EMBL:X15461

C/Superfamily: hepatitis A virus genome polyprotein

C/Keywords: coat protein; core protein; polyprotein

F/1-27/Product: coat protein 1A #status predicted <C1A>

F/28-495/Product: coat protein 1B #status predicted <C1B>

F/496-795/Product: coat protein 1C #status predicted <C1C>

F/796-984/Product: coat protein 1D #status predicted <C1D>

F/985-1091/Product: core protein 2A #status predicted <C2A>

F/1092-1426/Product: core protein 2B #status predicted <C2B>

F/1427-1499/Product: core protein 2C #status predicted <C2C>

F/1499-1521/Product: protein 3A #status predicted <P3A>

F/1522-1741/Product: protein 3B #status predicted <P3B>

F/1742-2230/Product: protein 3C #status predicted <P3C>

F/1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 92.0%; Score 911; DB 1; Length 2230;

Best Local Similarity 89.9%; Pred. No. 1.8e-72;

Matches 170; Conservative 14; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SHIECRKRYKELRLVGVGKORUKYAQELSNVLPFPRKMKGLFSQAKISLFTYTHEHIMK 60

Db 796 SHIECRKRYKELRLVGVGKORUKYAQELSNVLPFPRKMKGLFSQAKISLFTYTHEHIMK 855

Qy 61 FSWGVTADTALRFRGSLAAGRSVMTLEMDAGVLTGRLILNPKKINLADRMJGL 120

Db 856 LSWKGLTADTALRFRGSLAAGRSVMTLEMDAGVLTGRLILNPKKINLADRMJGL 915

Qy 121 EKFTSNKYKWSKVNPFHGMJLDEEIAANSKDPFNNSETDLCFLHMLNPKKINLADRMJGL 180

Db 916 EKFTSNKYKWSKVNPFHGMJLDEEIAANSKDPFNNSETDLCFLHMLNPKKINLADRMJGL 975
Qy 181 SGVOEIKEQ 189
Db 976 SGVOEIKEQ 984

RESULT 6

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)

N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein

C/Species: human hepatitis A virus

C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

C/Accession: A03905

R/Baxoudy, B.M.; Ticehurst, J.R.; Mele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinstrom

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A/Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA

A/Reference number: A03905; MUID:85166289; PMID:2984684

A/Accession: A03905

A/Molecule type: genomic RNA

A/Residues: 1-1358 <BAR>

C/Superfamily: hepatitis A virus genome polyprotein

C/Keywords: coat protein; core protein; polyprotein

F/1-245/Product: coat protein 1A #status predicted <C1A>

F/246-491/Product: coat protein 1B #status predicted <C1B>

F/492-836/Product: coat protein 1C #status predicted <C1C>

F/837-854/Product: core protein 2A (fragment) #status predicted <C2A>

F/855-1358/Product: core protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 33.6%; Score 333; DB 2; Length 1358;

Best Local Similarity 100.0%; Pred. No. 2.6e-21;

Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKRYKELRLVGVGKORUKYAQELSNVLPFPRKMKGLFSQAKISLFTYTHEHIMK 60

Db 792 SHIECRKRYKELRLVGVGKORUKYAQELSNVLPFPRKMKGLFSQAKISLFTYTHEHIMK 851

Qy 61 FSW 63

Db 852 FSW 854

RESULT 7

GENMYA

genome polyprotein - human hepatitis A virus (strain CR326) (fragment)

N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C/Species: human hepatitis A virus

A/Note: host Homo sapiens (man)

C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C/Accession: A03904

R/Innemyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W

J. Virol. 54, 247-255, 1985

A/Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A/Reference number: A03904; MUID:85185648; PMID:2985793

A/Accession: A03904

A/Molecule type: genomic RNA

A/Residues: 1-852 <LIN>

A/Cross-references: EMBL:M10033; NID:9329592; PIDN:AAA45470.1; PID:9329593

C/Superfamily: hepatitis A virus genome polyprotein

C/Keywords: coat protein; core protein; polyprotein

F/1-245/Product: coat protein 1A #status predicted <C1A>

F/246-491/Product: coat protein 1B #status predicted <C1B>

F/492-836/Product: coat protein 1C #status predicted <C1C>

F/837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 32.0%; Score 317; DB 1; Length 852;

Best Local Similarity 100.0%; Pred. No. 3.9e-20;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKRYKELRLVGVGKORUKYAQELSNVLPFPRKMKGLFSQAKISLFTYTHEHIMK 60

Db 792 SHIECRKRYKELRLVGVGKORUKYAQELSNVLPFPRKMKGLFSQAKISLFTYTHEHIMK 851

Qy 61 F 61
Db 852 F 852

RESULT 8

S04137

genome polypeptide - human hepatitis A virus (strain LCDC-1) (fragment)

C:Species: human hepatitis A virus

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000

C:Accession: S04137

R:Anderson, A.P.; Lau, P.C.K.; Chaudhary, R.

Nucleic Acids Res. 17, 3594, 1989

A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus

A:Reference number: S04137; MUID:89263805; PMID:2542903

A:Accession: S04137

A:Molecule type: mRNA

A:Residues: 1-341 <AND>

A:Cross-references: EMBL:X14666; NID:962301; PIDN:CAA32794.1; PID:94377576

A:Genetic:

A:Gene: VP1

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; polypeptide

F:2-340/Product: coat protein ID (VP1) #status predicted <MAT>

Query Match 23.8%; Score 236; DB 2; Length 341;

Best Local Similarity 97.8%; Pred. No. 2e-13;

Matches 45; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPYKELRLVGVKORLKYAOEELNEVLPPRKKMGKLFSSQA 46

Db 296 SHIECKPYKELRLVGVKORLKYAOEELNEVLPPRKKMGKLFSSQS 341

RESULT 9

GNV52

genome polypeptide - simian hepatitis A virus (strain CY-145) (fragment)

N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat protein

C:Species: simian hepatitis A virus

A>Note: host Macaca fascicularis (cynomolgus macaque)

C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999

C:Accession: J01180

R:Nathan, O.V.; Margolis, H.S.; Robertson, B.H.; Balaayan, M.; Brinton, M.A.

J. Gen. Virol. 72, 1685-1689, 1991

A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus mac

A:Reference number: J01180; MUID:91311421; PMID:1649902

A:Accession: J01180

A:Molecule type: genomic RNA

A:Residues: 1-839 <NAI>

A:Cross-references: GB:M59286; NID:9329599; PIDN:AAA45473.1; PID:9555083

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; glycoprotein; polypeptide

F:1-23/Product: coat protein 1A #status predicted <VP0>

F:24-45/Product: coat protein 1B #status predicted <VP3>

F:46-49/Product: coat protein 1C #status predicted <VP1>

F:492-839/Product: core protein 2A (fragment) #status predicted <P2>

F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.7%; Score 214.5; DB 1; Length 839;

Best Local Similarity 85.7%; Pred. No. 5.1e-11;

Matches 42; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

Qy 1 SHIECKPYKELRLVGVKORLKYAOEELNEVLPPRKKMGKLFSSQA 49

Db 792 SHIECKPYKELRLVGVKORLKYAOEELNEVLPPRKKMGKLFSSQS 839

RESULT 10

P00427

genome polypeptide - human hepatitis A virus (strain EP-35.730) (fragment)

C:Species: human hepatitis A virus

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: P00427

R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nathan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992

A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr

A:Reference number: P00427; MUID:92300330; PMID:1318940

A:Accession: P00427

A:Molecule type: mRNA

A:Residues: 1-56 <ROB>

C:Comment: This protein is from the VP1/2A junction region.

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; polypeptide

Query Match 14.5%; Score 144; DB 2; Length 56;
Best Local Similarity 96.4%; Pred. No. 3.2e-06;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPYKELRLVGVKORLKYAOEEL 28

Db 29 SHIECKPYKELRLVGVKORLKYAOEEL 56

RESULT 11

P00429

genome polypeptide - human hepatitis A virus (strain PRCl6) (fragment)

C:Species: human hepatitis A virus

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: P00429

R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nathan, O.V.; Siegl, G.; Widel

J. Gen. Virol. 73, 1365-1377, 1992

A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr

A:Reference number: P00427; MUID:92300330; PMID:1318940

A:Accession: P00429

A:Molecule type: mRNA

A:Residues: 1-56 <ROB>

A>Note: this protein is from the VP1/2A junction region

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; polypeptide

Query Match 14.4%; Score 143; DB 2; Length 56;
Best Local Similarity 96.4%; Pred. No. 4e-06;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPYKELRLVGVKORLKYAOEEL 28

Db 29 SHIECKPYKELRLVGVKORLKYAOEEL 56

RESULT 12

P00430

genome polypeptide - human hepatitis A virus (strain S23-1) (fragment)

C:Species: human hepatitis A virus

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: P00430

R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nathan, O.V.; Siegl, G.; Widel

J. Gen. Virol. 73, 1365-1377, 1992

A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr

A:Reference number: P00427; MUID:92300330; PMID:1318940

A:Accession: P00430

A:Molecule type: mRNA

A:Residues: 1-56 <ROB>

A>Note: this protein is from the VP1/2A junction region

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; polypeptide

Query Match 14.0%; Score 139; DB 2; Length 56;
Best Local Similarity 96.4%; Pred. No. 9e-06;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHIECKPYKELRLVGVKORLKYAOEEL 28

Db 29 SHIECKPYKELRLVGVKORLKYAOEEL 56

RESULT 13

P00427

genome polypeptide - human hepatitis A virus (strain EP-35.730) (fragment)

C:Species: human hepatitis A virus

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: P00427

Db 29 SHIECRKPYKELRLLEVGVKQRLKYAOEEL 56
Search completed: April 2, 2003, 11:11:26
Job time : 28.6189 secs

P00431
genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)
C/Species: human hepatitis A virus
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C/Accession: P00431
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuba, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A/Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A/Reference number: P00427; PMID:92300330; PMID:1318940
A/Accession: P00431
A/Molecule type: mRNA
A/Residues: 1-56 <ROB>
A/Note: this protein is from the VP1/2A Junction region
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; polyprotein

Query Match 13.8%; Score 137; DB 2; Length 56;
Best Local Similarity 96.4%; Pred. No. 1.4e-05;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 SHIECRKPYKELRLLEVGVKQRLKYAOEEL 28
Db 29 SHIECRKPYKELRLLEVGVKQRLKYAOEEL 56

RESULT 14

P00432
genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)
C/Species: human hepatitis A virus
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C/Accession: P00432
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuba, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A/Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A/Reference number: P00427; PMID:92300330; PMID:1318940
A/Accession: P00432
A/Molecule type: mRNA
A/Residues: 1-56 <ROB>
A/Note: this protein is from the VP1/2A Junction region
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; polyprotein

Query Match 13.7%; Score 136; DB 2; Length 56;
Best Local Similarity 96.4%; Pred. No. 1.7e-05;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 SHIECRKPYKELRLLEVGVKQRLKYAOEEL 28
29 SHIECRKPYKELRLLEVGVKQRLKYAOEEL 56

RESULT 15

P00428
genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C/Species: human hepatitis A virus
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C/Accession: P00428
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuba, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A/Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A/Reference number: P00427; PMID:92300330; PMID:1318940
A/Accession: P00428
A/Molecule type: mRNA
A/Residues: 1-56 <ROB>
A/Note: this protein is from the VP1/2A Junction region
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; polyprotein

Query Match 13.6%; Score 135; DB 2; Length 56;
Best Local Similarity 96.4%; Pred. No. 2.1e-05;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 SHIECRKPYKELRLLEVGVKQRLKYAOEEL 28

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 11:03:29 ; Search time 13.5345 Seconds
(without alignments)
579.187 Million cell updates/sec

Title: US-10-104-966-12_COPY_792_980
Perfect score: 990
Sequence: 1 SHIECKRKYKELRLEVGRQR.....KINLADRLGLSGVGEIKKQ 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
al number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	990	100.0	2227	1	P06617 hepatitis a
2	990	100.0	2227	1	P06441 hepatitis a
3	984	99.4	2227	1	P13901 hepatitis a
4	982	99.2	2226	1	P26581 hepatitis a
5	982	99.2	2226	1	P26582 hepatitis a
6	978	98.8	2226	1	P26580 hepatitis a
7	911	92.0	2230	1	P14553 simian hepa
8	317	32.0	852	1	P06442 hepatitis a
9	236	23.8	341	1	P13672 hepatitis a
10	214.5	21.7	839	1	P13788 simian hepa
11	188	19.0	808	1	P02381 hepatitis a
12	86	8.7	1289	1	MCE_REOVD
13	85	8.6	297	1	RM35_MOUSE
14	85	8.6	381	1	RIR2_CABEL
15	84.5	8.5	1353	1	CY99_HUMAN
16	80.5	8.1	1034	1	GCSP_YEAST
17	80	8.0	569	1	MALT_CANAL
18	79.5	8.0	355	1	TRMU_BORBU
19	79.5	8.0	1006	1	BAT1_YEAST
20	79.5	8.0	1250	1	SSD1_YEAST
21	79.5	8.0	1353	1	CY99_MOUSE
22	79.5	8.0	1374	1	RPOB_RICPR
23	78	7.9	739	1	RELA_STREO
24	78	7.9	1200	1	ATY2_MOUSE
25	78	7.9	1204	1	ATY2_HUMAN
26	78	7.9	1225	1	SMC1_YEAST
27	76.5	7.7	381	1	SWI_YEAST
28	76	7.7	581	1	DPO4_YEAST
29	75.5	7.6	297	1	ICIA_ECOLI
30	75.5	7.6	297	1	ICIA_SALTI
31	75.5	7.6	297	1	ICIA_SALTY
32	74	7.5	412	1	ALMT_CALCIN
33	74	7.5	842	1	AMPN_IACDL

34	73.5	7.4	481	1	APRF_PSEAE
35	73.5	7.4	1328	1	EXO2_SCHPO
36	73	7.4	897	1	APG1_YEAST
37	73	7.4	956	1	SVI_AQUAE
38	73	7.4	1099	1	PLCI_CANAL
39	73	7.4	1440	1	SYEP_HUMAN
40	72.5	7.3	660	1	ALIA_STPRN
41	72.5	7.3	971	1	US51_MOUSE
42	72.5	7.3	972	1	CY99_HUMAN
43	72.5	7.3	1305	1	CY99_XENLA
44	72	7.3	388	1	RECA_STPRN
45	72	7.3	863	1	ARS6_YEAST

ALIGNMENTS

RESULT 1

ID	POLG_HPAVH	STANDARD	PRT	2227 AA
AC	P06617; P06443; Q81082;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Genome polypeptide (Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)).			
DE	P3D (EC 2.7.7.48)).			
OS	Hepatitis A virus (strain HM-175).			
OC	Hepatitis A virus positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.			
OX	NCBI_TaxID=12098;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Wild type;			
RX	MEDLINE=87061253; PubMed=3023706;			
RA	Cohen J.T., Ticehurst J.R., Purcell R.H., Buckler-White A., Baroudy B.M.,			
RT	"Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses."			
RT	J Virol. 61:50-59(1987).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=Attenuated;			
RC	MEDLINE=87175701; PubMed=3031686;			
RX	Cohen J.T., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M., Purcell R.H.,			
RT	"Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with wild-type virus."			
RT	Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).			
RL	[3]			
RP	SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.			
RX	MEDLINE=85166289; PubMed=2984684;			
RA	Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr., Purcell R.H., Feinstone S.M.,			
RT	"Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA polymerase."			
RT	Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).			
RL	-I- CARAYUTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).			
CC	-I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.			
CC	-I- PMV: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.			
CC	-I- MICELLAROUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.			
CC	-I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.			
CC	-I- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT SHOWN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its			

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M4114; AAA45475.1; -
 CC EMBL; M41707; AAA45465.1; -
 CC EMBL; M41707; AAA45466.1; ALT_INIT.
 CC EMBL; M16632; AAA45471.1; -
 CC PIR; A25981; GNNYHM.
 CC PIR; A25914; GNNYMK.
 CC PIR; A03905; A03905.
 CC MEROPS; C03.005; -
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC RNA-directed RNA polymerase; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Thiol protease.
 CC CHAIN 1 23
 CC CHAIN 24 245
 CC CHAIN 246 491
 CC CHAIN 492 836
 CC CHAIN 837 980
 CC CHAIN 981 1087
 CC CHAIN 1088 1422
 CC CHAIN 1423 1496
 CC CHAIN 1497 1519
 CC CHAIN 1520 1738
 CC CHAIN 1739 2227
 CC CHAIN 77 77
 CC CHAIN 764 764
 CC CHAIN 821 821
 CC CHAIN 1052 1052
 CC CHAIN 1062 1062
 CC CHAIN 1118 1118
 CC CHAIN 1151 1151
 CC CHAIN 1163 1163
 CC CHAIN 1277 1277
 CC CHAIN 1500 1500
 CC CHAIN 1805 1805
 CC CHAIN 1930 1930
 CC CHAIN 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;
 CC SEQUENCE

Query Match
 Best Local Similarity 100.0%; Score 990; DB 1; Length 2227;
 Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SHIECRKPYKELRLVGVKQRLKYAQBELSNVLPFRKMKGLFSQAKISLFTYEEHEIMK 60
 792 SHIECRKPYKELRLVGVKQRLKYAQBELSNVLPFRKMKGLFSQAKISLFTYEEHEIMK 851

61 FSWRGVYADTRALRRFGFSILAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 120
 852 FSWRGVYADTRALRRFGFSILAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 911

121 EKTSTNKNYKSNFPGMDLBEIANSKDPFPMSETDLCFLHWNLPKKNINLADRMGL 180
 912 EKTSTNKNYKSNFPGMDLBEIANSKDPFPMSETDLCFLHWNLPKKNINLADRMGL 971

181 SGVOEIKEQ 189
 972 SGVOEIKEQ 980

RESULT 2
 POLG_HPAVL STANDARD; PRT; 2227 AA.
 ID POLG_HPAVL
 AC P06471;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48).
 OS Hepatitis A virus (strain LA).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 CC NCBI_TaxID=12099;
 CC (1)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=85190549; PubMed=2986127;
 RA Najarian R., Caput D., Gee W.W., Porter S.J., Renard A.,
 RA Merryweather J., van Neer G., Dina D.,
 RT "Primary structure and gene organization of human hepatitis A virus";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA} (N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; K02990; AAA45472.1; -
 CC PIR; A03903; GNNYHR.
 CC MEROPS; C03.005; -
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC RNA-directed RNA polymerase; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Thiol protease.
 CC CHAIN 1 23
 CC CHAIN 24 245
 CC CHAIN 246 491
 CC CHAIN 492 836
 CC CHAIN 837 980
 CC CHAIN 981 1076
 CC CHAIN 1077 1484
 CC CHAIN 1423 1484
 CC CHAIN 1485 1507
 CC CHAIN 1508 1678
 CC CHAIN 1679 2227
 CC CHAIN 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;
 CC SEQUENCE

Query Match
 Best Local Similarity 100.0%; Score 990; DB 1; Length 2227;
 Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SHIECRKPYKELRLVGVKQRLKYAQBELSNVLPFRKMKGLFSQAKISLFTYEEHEIMK 60
 792 SHIECRKPYKELRLVGVKQRLKYAQBELSNVLPFRKMKGLFSQAKISLFTYEEHEIMK 851

61 FSWRGVYADTRALRRFGFSILAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 120
 852 FSWRGVYADTRALRRFGFSILAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKIVSLI 911

121 EKTSTNKNYKSNFPGMDLBEIANSKDPFPMSETDLCFLHWNLPKKNINLADRMGL 180
 912 EKTSTNKNYKSNFPGMDLBEIANSKDPFPMSETDLCFLHWNLPKKNINLADRMGL 971

181 SGVOEIKEQ 189
 972 SGVOEIKEQ 980

RESULT 3

POLG_HPAVM STANDARD: PRT: 2227 AA
 AC P13901: Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
 AC Q81090; Q81091; Q81092; Q81093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
 P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain MBB).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
 Deinhardt F.;
 RT "The entire nucleotide sequence of the genome of human hepatitis A
 virus (isolate MBB)."
 CC Virus Res. 8:153-171(1987).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 {RNA} (N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
 WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M20273; AAA45474.1; -.
 DR PIR: J50303; GNNYHB.
 DR MEROPS: C03.005; -.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 KM Polypeptide; Coat protein; Core protein; Transferrase;
 KM RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 FT CHAIN 2227 AA; 251425 MW; EC983ED2A7C86349 CRG64;
 SQ SEQUENCE
 Query Match 99.4%; Score 984; DB 1; Length 2227;
 Best Local Similarity 99.5%; Pred. No. 2.2e-83;
 Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	121	EXFTNKWKSKNPKPHGMLDLEIANSKDPNNMETDLCFLHLWNPKNKMLAARMGL	180
Db	912	EXFTNKWKSKNPKPHGMLDLEIANSKDPNNMETDLCFLHLWNPKNKMLAARMGL	971
Qy	181	SGVQEIKEQ 189	
Db	972	SGVQEIKEQ 980	
RESULT 4			
ID	POLG_HPVA4	STANDARD;	PRT; 2226 AA.
AC	P26581;		
DT	01-AUG-1992	(Rel. 23, Created)	
DT	01-AUG-1992	(Rel. 23, Last sequence update)	
DT	15-JUN-2002	(Rel. 41, Last annotation update)	
DE	Genome polypeptide [Contants: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]		
DE	P3D (EC 2.7.7.48)]		
DE	Hepatitis A virus (strain 43c)		
OC	Virusess; ssRNA positive-strand virusess, no DNA stage; Picornaviridae; Hepatovirus.		
OX	NCBI_TaxID=12095;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91162758; PubMed=1705995;		
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Felnstone S.M., Cromeans T., Jansen R.W.;		
RT	"Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."		
RL	J. Virol. 65:2056-2065(1991).		
CC	-I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).		
CC	-I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.		
CC	-I- PMU: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.		
CC	-I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; M59809; AAA5469.1; -.		
DR	MEROPS; C03.005; -.		
DR	InterPro; IPR001205; RNA_helicase.		
DR	InterPro; IPR001205; RNA_pol_P3D.		
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.		
DR	Pfam; PF00910; RNA_helicase; 1.		
KW	Polypeptide; Coat protein; Core protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.		
KW	CHAIN 1 23		
FT	CHAIN 24 245	COAT PROTEIN VP4 (P1A) .	
FT	CHAIN 246 491	COAT PROTEIN VP2 (P1B) .	
FT	CHAIN 492 794	COAT PROTEIN VP3 (P1C) .	
FT	CHAIN 795 900	COAT PROTEIN VP1 (P1D) .	
FT	CHAIN 901 1087	CORE PROTEIN P2A.	
FT	CHAIN 1088 1422	CORE PROTEIN P2B.	
FT	CHAIN 1423 1495	CORE PROTEIN P2C.	
FT	CHAIN 1496 1518	PROBABLE PROTEIN P3A.	
FT	CHAIN 1519 1737	PROBABLE PROTEIN P3B.	
FT	CHAIN 1738 2226	PROBABLE PROTEIN P3C.	
FT	CHAIN 1738 2226	RNA-DIRECTED POLYMERASE P3D.	
SO	SEQUENCE 2226 AA; 251107 MW; 403B4CA80809BF75 CRC64;		
Query Match	99.2%;	Score 982;	DB 1; Length 2226;
Best local Similarity	98.9%;	Pred. No. 3.4e-83;	
Matches 187;	Conservative 1;	Mismatches 1;	Indels 0; Gaps 0;

```

OY 1 SHIECRKPKYKELRLVEGKORLKYAOEELSNEVLPPPRKMGKLSFOAKISLFYTEHEHMK 60
DB 792 SHIECRKPKYKELRLVEGKORLKYAOEELSNEVLPPPRKMGKLSFOAKISLFYTEHEHMK 851
OY 61 FSWRGVTDADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTDADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
OY 121 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 971
OY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 5
POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 18f).
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59808; AAA45467.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR Pfam; PF00680; RNA_dep_RNA_pol.1.
DR Pfam; PF00910; RNA_helicase.1.
KW Polypeptide; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
KW PROBABLE PROTEIN P3A.

```

```

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D68 CRC64;

Query Match 99.2%; Score 982; DB 1; Length 2226;
Best Local Similarity 98.9%; Pred. No. 3,4e-83;
Matches 187; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLVEGKORLKYAOEELSNEVLPPPRKMGKLSFOAKISLFYTEHEHMK 60
DB 792 SHIECRKPKYKELRLVEGKORLKYAOEELSNEVLPPPRKMGKLSFOAKISLFYTEHEHMK 851
OY 61 FSWRGVTDADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 120
DB 852 FSWRGVTDADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLRLNDEKWTMCKDKIVSLI 911
OY 121 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 971
OY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 6
POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 24a).
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59810; AAA45468.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol.1.
DR Pfam; PF00910; RNA_helicase.1.
KW Polypeptide; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.

```

FT	CHAIN	1	23	COAT PROTEIN VP4 (PIA).
FT	CHAIN	24	245	COAT PROTEIN VP2 (PIB).
FT	CHAIN	246	491	COAT PROTEIN VP3 (PIC).
FT	CHAIN	492	794	COAT PROTEIN VP1 (PID).
FT	CHAIN	795	500	CORE PROTEIN P2A.
FT	CHAIN	501	1087	CORE PROTEIN P2B.
FT	CHAIN	1088	1422	CORE PROTEIN P2C.
FT	CHAIN	1423	1495.	PROBABLE PROTEIN P3A.
FT	CHAIN	1496	1518	PROBABLE PROTEIN P3B.
FT	CHAIN	1519	1737	PROBABLE PROTEIN P3C.
FT	CHAIN	1738	2226	RNA-DIRECTED POLYMERASE P3D
50	SEQUENCE	2226 AA;	251152 MW;	6CDB5A91D654E2BF CRC64;

Query Match	98.8%	Score 978;	DB 1;	Length 2226;
Best Local Similarity	98.4%	Pred. No. 8.1e-83;		
Matches 186;	Conservative	2;	Mismatches 1;	Indels 0;
			Gaps	0;

Qy	1	SHBEKRYKXELRLVEVGQRILKVAQOELSNVLEPPPRKMGLEFQAKISLFLPYEEHIMK	60
Dh	792	SHIEKRYKXELRLVEVGQRILKVAQOELSNVLEPPPRKMGLEFQAKISLFLPYEEHIMK	851
Dh	61	FSMWGTADTALRALRFGFSLAAGRSWLTENDAGVLIGRLIRLNDKXTEKODKDIYSLI	120
Dh	852	FSMWGTADTALRALRFGFSLAAGRSWLTENDAGVLIGRLIRLNDKXTEKODKDIYSLI	911
Qy	121	EKFLSNKRWKSKVNPFGHGLDLBEELIANSKDFPPNNSETDLCFLHMLNPKKINLADRMGL	180
Dh	912	EKFLSNKRWKSKVNPFGHGLDLBEELIANSKDFPPNNSETDLCFLHMLNPKKINLADRMGL	971
Qy	181	SGVOEIKEQ	189
Dh	972	SGVOEIKEQ	980

RESULT 7
POLG_HPAYS
ID_POLG_HPAYS STANDARD; PRT; 2230 AA

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptin [Comments: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DS Simian hepatitis A virus (strain AGM-27).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae.
OC Hepatovirus.
NC NCB1_Taxid=12102;

SEQUENCE FROM N.A.
MEDLINE=91311420; Pubmed=1649901;
RA Tseurev S.A., Emerson S.U., Balaayan M.S., Tichenurst J.R.,
Purcell R.H.,
RT "Stimian hepatitis A virus (HAV) strain AGM-27: comparison of genome
structure and growth in cell culture with other HAV strains.";
RU J. Gen. Virol. 72:1677-1683(1991).

RP
SEQUENCE OF 1750-2164 FROM N.A.
RX
MEDLINE=89232168; Pubmed=2541023;
RA
Batalvan M.S., Kusov Y.Y., Andjardzidze A.G., Tsarev S.A.;
RA
Savaldov E.D., Chizhikov V.E., Bairov V.M., Vasilchenko S.K.;
RT
"variations in genome fragments coding for RNA polymerase in human
and simian hepatitis A viruses".
PL
FEBS Lett. 247:425-428 (1989).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC (RNA) (N).
CC
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2
CC VP3 AND VP4.

CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce> or send an email to license@isb-sib.ch).

DR	EMBL, D00322, EMBO168.1, -	
DR	EMBL, X15461, CA33490.1, -	
DR	PIR, A30470, GNNYSA.	
DR	PIR, S04885, S04885.	
DR	MEROPS, C03.005, -	
DR	InterPro, IPR000605, RNA_helicase.	
DR	InterPro, IPR001205, RNA_pol_P3D.	
DR	Pfam, PF00680, RNA_dep_RNA_pol, 1.	
DR	Pfam, PF00910, RNA_helicase, 1.	
KW	Polypeptide; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolyase; Thiol protease.	
KW	RNA-directed RNA polymerase; Hydrolyase; Thiol protease.	
FT	CHAIN 1 27	
FT	CHAIN 28 249	COAT PROTEIN VP4 (PI4) .
FT	CHAIN 250 495	COAT PROTEIN VP2 (PI2) .
FT	CHAIN 496 795	COAT PROTEIN VP3 (PI3) .
FT	CHAIN 796 984	COAT PROTEIN VP1 (PI1) .
FT	CHAIN 985 1091	CORE PROTEIN P2A.
FT	CHAIN 1092 1426	CORE PROTEIN P2B.
FT	CHAIN 1427 1498	CORE PROTEIN P2C.
FT	CHAIN 1499 1521	PROBABLE PROTEIN 3A.
FT	CHAIN 1522 1741	PROBABLE PROTEIN 3B.
FT	CHAIN 1742 2230	PROBABLE PROTEIN 3C.
FT	CHAIN 1742 2230	RNA-DIRECTED POLYMERASE 3D.
SO	SEQUENCE 2230 AA; 251296 MW; 87B3230E334E1F19 CRC64;	

Query Match	92.0%;	Score 911;	DB 1;	Length 2230;
Best Local Similarity	89.9%;	Pred. NO. 1.4e-76;		
Matches 170;	Conservative 14;	Mismatches 5;	Indels 0;	Gaps 0;

Qy	1	SHIEGRKRYKYLRLVEVGKORLKYAEOEELSNVLEPPPRKMGFLSQOAKISLPTYTEHSHIMK	60
Dd	796	SHIEGRKRYKYLRLVEVGKORLKYAMEELSNELPPPRKMGFLSQOAKISLPTYTEHSHIVK	855
Qy	61	FSMRGVTADTALRLRFGFSLAAGHVSVTLLEMDACVLTGRLLRLNDEKXTEKMDKIVSLI	120
Dd	856	LSWKGTLTADTALRLRRYGFSLAAGHVSVTLLEMDACVLTGRMLRLNDEKXTEIKODKIVLALV	915
Qy	121	EKFSTSNKYSKYNFPHGMLDLEELIAANSKDFPNNSSEFDLCFLHLWLNPKKINLADRMGL	180
Dd	916	EKFSTSNKYSKYNFPHGMLDLEELIASNSKDFPNNSSEFDLCFLHLWLNPKKINLADRMGL	975
Qy	181	SGVOEIKEO	189
Dd	976	SGVOEIKEO	984

```

RESULT 8
POLG_HPAVC          STANDARD;      PRT;      852 AA
ID_POLG_HPAVC
AC P06442; Q83741; Q83742;

```

DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptin [Contents: Coat proteins VP1 To VP4; Core protein p21 (Fragment)].
DE Hepatitis A virus (strain CR326).
OS viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae.
OC Hepatovirus.
NCBI_TaxID=12097;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85185648; PubMed=2985793;
 RA Linemeyer D.L., Menze J.G., Martin-Gallardo A., Hughes J.V.,
 w/ J. Ostr.

RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA."
 RL J. Virol. 54:247-255(1985).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,

CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M10033; AAA45470.1; -.
CC PIR: A03904; GNNYHA.
CC
CC Polypeptide: Coat protein; Core protein.
CC
CC FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
CC FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
CC FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
CC FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
CC FT CHAIN 837 >852 CORE PROTEIN P2A.
CC FT NON TER 852 852
CC SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;
CC
CC Query Match 32.0%; Score 317; DB 1; Length 852;
CC Best Local Similarity 100.0%; Pred. No. 9.5e-22;
CC Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 SHIECKRKYKELRLEVKGKRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYEEHEIMK 60
CC DB 792 SHIECKRKYKELRLEVKGKRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYEEHEIMK 851
CC
CC QY 61 F 61
CC DB 852 F 852
CC
CC RESULT 9
CC POLG_HPAV1 STANDARD; PRT; 341 AA.
CC ID POLG_HPAV1 STANDARD; PRT; 341 AA.
CC AC P13672;
CC DT 01-JAN-1990 (Rel. 13, Created)
CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Genome polypeptide [contains: Coat proteins VP1 TO VP3; Core protein
CC P2A] (Fragment).
CC OS Hepatitis A virus (strain LCD-1).
CC VS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_Taxid=12093;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=89263805; PubMed=2542903;
CC RA Andonov A.P., Lau P., Chaudhary R.;
CC RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
CC hepatitis A virus (HAV)."
CC RL Nucleic Acids Res. 17:3594-3594(1989).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X14666; CAA32794.1; -.
CC PIR: S04137; S04137.
CC Polypeptide: Coat protein; Core protein.

FT NON TER 1 1
FT CHAIN <1 1 COAT PROTEIN VP3 (1C).
FT CHAIN 2 340 COAT PROTEIN VP1 (1D).
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289B12655 CRC64;
CC
CC Query Match 23.8%; Score 236; DB 1; Length 341;
CC Best Local Similarity 97.8%; Pred. No. 1.1e-14;
CC Matches 45; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 SHIECKRKYKELRLEVKGKRLKYAOEELSNEVLPPPRKMKGLFSQAKIS 46
CC DB 236 SHIECKRKYKELRLEVKGKRLKYAOEELSNEVLPPPRKMKGLFSQAKIS 341
CC
CC RESULT 10
CC POLG_HPAV1 STANDARD; PRT; 839 AA.
CC ID POLG_HPAV1 STANDARD; PRT; 839 AA.
CC AC P31788;
CC DT 01-JUL-1993 (Rel. 26, Created)
CC DT 01-JUL-1993 (Rel. 26, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core protein
CC P2A] (Fragment).
CC OS Simian hepatitis A virus (strain CY-145).
CC VS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_Taxid=31707;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=9131421; PubMed=1649902;
CC RA Naiman O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
CC "Sequence analysis of a new hepatitis A virus naturally infecting
CC cynomolgus macaques (Macaca fascicularis)."
CC RT J. Gen. Virol. 72:1685-1689(1991).
CC RL J. Gen. Virol. 72:1685-1689(1991).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59286; AAA45473.1; -.
CC PIR: J01180; GNNYS2.
CC RN Polypeptide: Coat protein; Core protein.
CC KM Polypeptide: Coat protein; Core protein.
CC FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
CC FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
CC FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
CC FT CHAIN 492 ? COAT PROTEIN VP1 (P1D).
CC FT CHAIN ? >839 CORE PROTEIN P2A.
CC FT NON TER 839 839
CC SQ SEQUENCE 839 AA; 93825 MW; 2CACC4BD1E192DBC CRC64;
CC
CC Query Match 21.7%; Score 214.5; DB 1; Length 839;
CC Best Local Similarity 85.7%; Pred. No. 3.3e-12;
CC Matches 42; Conservative 5; Mismatches 1; Indels 1; Gaps 1;
CC
CC QY 1 SHIECKRKYKELRLEVKGKRLKYAOEELSNEVLPPPRKMKGLFSQAKIS 49
CC DB 792 SHIECKRKYKELRLEVKGKRLKYAOEELSNEVLPPPRKMKGLFSQAKIS 839
CC
CC RESULT 11
CC POLG_HPAV1 STANDARD; PRT; 808 AA.
CC ID POLG_HPAV1 STANDARD; PRT; 808 AA.
CC AC Q02381;

RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Watanabe-Borja A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 CC Nature 409:685-690(2001).
 CC -1- SUCCINYL-L-GLUTAMATE SYNTHETASE.
 CC -1- SIMILARITY: TO THE N-TERMINAL OF THREONYL-TRNA SYNTHETASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF239728; AAF44697.1; -
 CC EMBL; BC012274; AAH12274.1; -
 CC EMBL; BC016561; AAH16561.1; -
 CC EMBL; AK007681; BAB25185.1; -
 CC MGD; WGI:J351620; MRP139.
 CC InterPro; IPR004095; TGS_dom.
 CC Pfam; PF02824; TGS_1.
 CC Ribosomal protein; Mitochondrion.
 CC FT CONFLICT 94 94 K -> R (IN REF. 2; AAH16561).
 CC FT CONFLICT 125 125 L -> I (IN REF. 2; AAH16561).
 CC FT CONFLICT 244 244 L -> F (IN REF. 1).
 CC FT CONFLICT 276 278 ERF -> GRS (IN REF. 2; AAH16561).
 CC FT CONFLICT 278 278 F -> S (IN REF. 2; AAH12274 AND 3).
 CC SEQUENCE 297 AA; 34559 MW; 7E0461EBCB875C57F CRC64;
 Query Match 8.6%; Score 85; DB 1; Length 297;
 Best Local Similarity 27.8%; Pred. No. 1.1;
 Matches 42; Conservative 14; Mismatches 37; Indels 58; Gaps 8;
 QY 60 KESWGVATDTRAL--RRQFSLAAGRSVWTLMDAGVLTGRLRLNDEKXTEMDKDIIV 117
 DB 154 KENLRSTFTDAHALIYRDLDFE-----TLVDVAV----- 183
 QY 118 SLIEKTSNKYKSVNF-----PHGMLDEEIIANSKDPNNMSE-----TDLCF-- 161
 DB 184 -ALEIQTGHKKY--KVPFIEEKASQNPRIYKLRIG--DFIDVSEGLPIPTSCFOQ 236
 QY 162 ---LHMLNPKKINLADRMGLSGVGEIKQ 189
 DB 237 EVSAVHNLPNSQPNLIRFGSLPFTHLRAQ 267
 RESULT 14
 RIR2 CAEEL
 ID RIR2 CAEEL STANDARD; PRT; 381 AA.
 AC P42170;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
 DE (Ribonucleoside reductase).
 GN RNR-2 OR CO3C10.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OK NCBI_Taxid=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Submitted (AUG-1994) to the EMBL/GenBank/DBI databases.
 CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
 CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
 CC thioredoxin + H(2O) = ribonucleoside diphosphate + reduced
 CC thioredoxin.

CC -1- COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).
 CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
 CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
 CC SMALL CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z35637; CAAB4688.1; -
 CC HSSP; P1157; IXSM.
 CC DR HSSP; P1157; IXSM.
 CC DR WormPep; CO3C10.3; CE00874.
 CC DR InterPro; IPR000358; Ribonuc1_reductase.
 CC DR Pfam; PF00268; ribonuc_red_sm; 1.
 CC DR PROSITE; PS00368; RIBORED_SMALL; 1.
 CC KM Oxidoreductase; DNA replication; Iron.
 CC FT METAL 130 130 IRON 1 (BY SIMILARITY).
 CC FT METAL 161 161 IRON 1 AND 2 (BY SIMILARITY).
 CC FT METAL 164 164 IRON 1 (BY SIMILARITY).
 CC FT METAL 224 224 IRON 2 (BY SIMILARITY).
 CC FT METAL 258 258 IRON 2 (BY SIMILARITY).
 CC FT METAL 261 261 IRON 2 (BY SIMILARITY).
 CC FT ACT SITE 168 168 BY SIMILARITY.
 CC SEQUENCE 381 AA; 44289 MW; 75497147ABF36C9 CRC64;
 Query Match 8.6%; Score 85; DB 1; Length 381;
 Best Local Similarity 21.0%; Pred. No. 1.5;
 Matches 48; Conservative 39; Mismatches 58; Indels 84; Gaps 12;
 QY 10 KKL-RLEVGKQRLKAOEELSNEVLPERRKMGLEFSQ-----KSLPYTEHEIMK 60
 DB 30 KLEKLEIIVDQTKASASSETNNE-----SEVNELDADPEMDLIDNRVIPPPLKHDHWN 84
 QY 61 PSWGVATDTRALRRFGSLAAGRSVWTL--EMDAGVLTGRLRLNDEKXTEMDKDIIV 115
 DB 85 FYKKAVA-----SFTVEEYDLK-----DNND--WEKNGDEQYFI 119
 QY 116 -----IV--SLIEKTSNKYKSVNFPHG--MLDLEEIIANSKDPNNMSEDTL-- 159
 DB 120 SRLNFPASDGVINENICERSNEVOVSARFVGFQIAIENI--HSEWYSKLIETIYIR 177
 QY 160 -----CFLHMLNPKKINLADRMGLSGVGEI 186
 DB 178 DETERNTLFNAVDEFEFTKKQADWALRWISDKKASFAERLIAFAAVEGI 226
 RESULT 15
 CYA9 HUMAN
 ID CYA9 HUMAN STANDARD; PRT; 1353 AA.
 AC O60503; Q9UGP2; O60273; Q9BWT4;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adenylate cyclase, type IX (EC 4.6.1.1) (ATP pyrophosphate-lyase)
 DE (Adenylate cyclase).
 GN ADENYLYL CYCLASE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OK NCBI_Taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98292499; PubMed=9628827;
 RA Hacker B.M., Tomlinson J.E., Wayman G.A., Sultana R., Chan G.,
 RA Villacres E., Distelche C., Storm D.R.;
 RT "Cloning, chromosomal mapping, and regulatory properties of the human
 RT type 9 adenylate cyclase (ADCY9).";

RL Genomics 50:97-104(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Paterson J.M., Smith S.M., Simpson J., Grace O.C., Bell J.E.,
 RA Antoni F.A.;
 RT "Cloning and characterisation of human adenylyl cyclase IX:
 RT differential mRNA regulation and inhibition by Ca^{2+} /calciocneurin.";
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Toyota T., Yamada K., Meerabux J., Hattori E., Saito K.,
 RA Yoshitsugu K., Shimizu H., Nankai M., Toru M., Decera-Wadleigh S.D.,
 RA Yoshikawa T.;
 RT "Mutation screening, case control study and transmission
 RT disequilibrium analysis of adenylyl cyclase 9 (ADCY9) gene in
 RT functional psychoses.";
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 788-1353 FROM N.A.
 RN TISSUE=Brain;
 MEDLINE=98290545; PubMed=9628581;
 Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 Nomura N., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 CC -1- FUNCTION: May play a fundamental role in situations where fine
 CC interplay between intracellular Ca^{2+} and cAMP determines the
 CC cellular function. May be a physiologically relevant docking site
 CC for calcineurin (by similarity).
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -1- ENZYME REGULATION: Insensitive to Ca^{2+} /calmodulin, forskolin and
 CC bombactonin. Stimulated by beta-adrenergic receptor activation.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
 CC FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 FKBP-TYPE PPIASE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 CC EMBL; AF036927; AAC24201.1; -
 CC EMBL; AJ133123; CAB65084.1; -
 CC EMBL; AY028959; AAK29464.1; -
 CC EMBL; AY028949; AAK29464.1; JOINED.
 CC EMBL; AY028950; AAK29464.1; JOINED.
 CC EMBL; AY028951; AAK29464.1; JOINED.
 CC EMBL; AY028952; AAK29464.1; JOINED.
 CC EMBL; AY028953; AAK29464.1; JOINED.
 CC EMBL; AY028954; AAK29464.1; JOINED.
 CC EMBL; AY028955; AAK29464.1; JOINED.
 CC EMBL; AY028956; AAK29464.1; JOINED.
 CC EMBL; AY028957; AAK29464.1; JOINED.
 CC EMBL; AB011092; BAA25446.1; -
 CC HSSP; P26769; IAB8.
 CC DR HSP; HGNC:240; ADCY9.
 CC DR MIM; 603302; -
 CC DR InterPro: IPR001054; G_cyclase.
 CC DR Pfam; PF00211; guanylate_cyc; 2.
 CC DR SMART; SM00044; CYC; 2.
 CC DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
 CC DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 2.
 CC DR Lyase; CAMP synthesis; Transmembrane; Glycoprotein; Repeat;
 CC isomerase; Rotamase.
 CC FT DOMAIN 1 117 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 118 138 POTENTIAL.
 CC FT

FT DOMAIN 139 141 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 142 162 POTENTIAL.
 FT DOMAIN 163 171 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 172 192 POTENTIAL.
 FT DOMAIN 193 215 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 216 235 POTENTIAL.
 FT DOMAIN 236 241 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 242 259 POTENTIAL.
 FT DOMAIN 260 280 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 281 301 POTENTIAL.
 FT DOMAIN 302 786 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 787 807 POTENTIAL.
 FT DOMAIN 808 818 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 819 839 POTENTIAL.
 FT DOMAIN 840 867 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 868 888 POTENTIAL.
 FT DOMAIN 889 891 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 892 912 POTENTIAL.
 FT DOMAIN 913 920 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 921 941 POTENTIAL.
 FT DOMAIN 942 975 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 976 996 POTENTIAL.
 FT DOMAIN 997 1353 POTENTIAL.
 FT CARBOHYD 504 611 PPIASE, FKBP-TYPE.
 FT TRANSMEM 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 493 493 R -> G (IN REF. 1 AND 3).
 FT CONFLICT 884 884 V -> A (IN REF. 2).
 FT CONFLICT 1154 1154 N -> S (IN REF. 2).
 FT CONFLICT 1253 1353 OHOLISPDIVQVDSIGRSPTEIANLVPSVOYDKTSL
 GDSSTQAKDHLRSKRPKWSPVKAEEGRGKALEKDDCD
 ETGIEANETLRLNVSKEV -> APAVHLPRHPRGGQOHR
 TVSHRDCQPGAPFCVCGDISGF (IN REF. 1).
 FT CONFLICT 1308 1308 R -> P (IN REF. 3 AND 4).
 FT SEQUENCE 1353 AA; 150858 MW; 4F43DA12690A9DBC CRC64;
 SQ
 Query Match 8.5%; Score 84.5; DB 1; Length 1353;
 Best Local Similarity 23.5%; Pred. No. 7.6;
 Matches 54; Conservative 30; Mismatches 83; Indels 63; Gaps 12;
 QY 3 IECKRPYELRL-----VGKQRLKIVAOEELSNEY-----LPPRRKMKGLFSQAKISL 50
 DB 323 LEVERALKERMIHVSVPRIADDLMKQGESESVKSHATSSPNRKXKSSIOKAPIAF 382
 QY 51 -FTYEEHEIMKFSRGVYADTRALRRFGSLAAGRSWTLEMDAGVLTGRILRLNDEKM 108
 DB 383 RPFKMOQIEVSIIFADIVGFTK-----MSANKSAHALVGLINDLFGDRDLCE--- 431
 QY 109 TEMKDKIVSLIEKFTSNKWSKVNFPH-----GMLD-LEELIANSKDPFN 153
 DB 432 -ETKCEKISTL-----GDCYVAGCPREPRADHAYCCIEEMGLGMKAIKAEQFOCKEKEMVN 485
 QY 154 M-----SETDLCFLH-----WLNPKKINLADRM--LGISGVQOEIKK 188
 DB 486 MRGVGHTRTVLGGLIGMRPFKDVSN--DVVLANLMEQLGVAGVHISE 533

Search completed: April 2, 2003, 11:08:34
 Job time : 17.5345 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2003, 11:04:04 ; Search time 51.7212 Seconds

(without alignments)
752.939 Million cell updates/sec

Title: US-10-104-966-12_COPY_792_980

Perfect score: 990
Sequence: 1 SHIECRKPKYKELRLVGVKOR.....KINLADRMGLGVQEIKEQ 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	990	100.0	1124	12	084780 hepatitis a
2	990	100.0	1161	12	005794 hepatitis a
3	990	100.0	2225	12	09DL32 hepatitis a
4	987	99.7	2227	12	09IFH5 hepatitis a
5	984	99.4	2216	12	09WMA2 hepatitis a
6	984	99.4	2218	12	067824 hepatitis a
7	984	99.4	2227	12	067825 hepatitis a
8	984	99.4	2227	12	067826 hepatitis a
9	984	99.4	2227	12	09WMA3 hepatitis a
10	981	99.1	2227	12	09WMA0 hepatitis a
11	978	98.8	2227	12	09WMA9 hepatitis a
12	978	98.8	2227	12	08QV03 hepatitis a
13	977	98.7	2227	12	09WMA1 hepatitis a
14	975	98.5	2218	12	067817 hepatitis a
15	973	98.3	2227	12	09WMA4 hepatitis a
16	970	98.0	2227	12	08V0N6 hepatitis a

17	947	95.7	2227	12	08QRI6 hepatitis a
18	914	92.3	184	12	087092 simian hepa
19	752	76.0	251	12	09ENQ9 hepatitis a
20	752	76.0	251	12	09ENQ6 hepatitis a
21	752	76.0	251	12	09ENQ5 hepatitis a
22	752	76.0	251	12	09ENQ4 hepatitis a
23	752	76.0	251	12	09ENQ7 hepatitis a
24	752	76.0	251	12	09ENQ2 hepatitis a
25	752	76.0	251	12	09ENQ6 hepatitis a
26	752	76.0	251	12	09ENQ4 hepatitis a
27	752	76.0	251	12	09ENQ2 hepatitis a
28	751	75.9	251	12	09ENQ7 hepatitis a
29	748	75.6	251	12	09ENQ5 hepatitis a
30	747	75.5	251	12	09ENQ1 hepatitis a
31	746	75.4	251	12	09ENQ8 hepatitis a
32	746	75.4	251	12	09ENQ3 hepatitis a
33	746	75.4	251	12	09ENQ8 hepatitis a
34	746	75.4	251	12	09ENQ3 hepatitis a
35	745	75.3	251	12	09ENQ1 hepatitis a
36	743	75.1	251	12	09ENQ2 hepatitis a
37	742	74.9	251	12	09ENQ0 hepatitis a
38	741	74.8	251	12	09ENQ4 hepatitis a
39	741	74.8	251	12	09ENQ9 hepatitis a
40	741	74.8	251	12	09ENQ8 hepatitis a
41	740	74.7	251	12	09ENQ9 hepatitis a
42	738	74.5	251	12	09ENQ5 hepatitis a
43	737	74.4	251	12	09ENQ0 hepatitis a
44	737	74.4	251	12	09ENQ6 hepatitis a
45	733	74.0	251	12	09ENQ0 hepatitis a

ALIGNMENTS

RESULT 1

ID 084780 PRELIMINARY; PRT; 1124 AA.

AC 084780;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE RNA for capsid VP4-VP1 and NS-proteins (NON-structural proteins)

DE (Fragment).

OC Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OC NCBI_taxid=12092;

RA SEQUENCE FROM N.A.

RA Oshimnikov Y.A., Sverdlov E.D., Tsarev S.A., Arsenyan S.G.,

RA Rokhlina T.O., Chizhikov V.E., Petrov N.A., Prikhod'ko G.G.,

RA Bilnov V.M., Vasilenko S.K., Sandakchiev L.S., Kusov Y.Y.,

RA Grabko V.I., Flier G.P., Baiyan M.S., Drozdov S.G.;

RL EMBL; X04200; CAA27797.1; -

DR EMBL; A11312; CAA00953.1; -

KW Nonstructural protein.

FT NON_TER 1 1124

FT NON_TER 1 1124

SO SEQUENCE 1124 AA; 127026 MW; 38449E2D2ABDF8CA CRC64;

Query Match 100.0%; Score 990; DB 12; Length 1124;

Best Local Similarity 100.0%; Pred. No. 4e-86;

Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SHIECRKPKYKELRLVGVKORLKYAQEELSNVLPKPKKGI.FSQAKISLPYTEHEIMK 60
DB	749	SHIECRKPKYKELRLVGVKORLKYAQEELSNVLPKPKKGI.FSQAKISLPYTEHEIMK 808
QY	61	FSWRGVTDTRALRFGFSLAAGRSVMTLMDAGVLTGRLIRLNDEKXTKEMDDKIVSLI 120
DB	809	FSWRGVTDTRALRFGFSLAAGRSVMTLMDAGVLTGRLIRLNDEKXTKEMDDKIVSLI 868

QY 121 EKFSTNKKYKSKVNFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMGL 180
 DB 869 EKFSTNKKYKSKVNFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMGL 928
 QY 181 SGVOEIKEQ 189
 DB 929 SGVOEIKEQ 937

RESULT 2

QY 005794 PRELIMINARY; PRT; 1161 AA.
 AC Q05794; Q67800; Q67801; Q67802; Q67803; Q67804; Q67805; Q67806;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Genome polyprotein (Coat proteins VP1 to VP4; core proteins P2A to P2C; probable proteins P3A to P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)) (Fragment).
 DE Hepatitis A virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 CC NCB1_TaxID=12092;
 RN SEQUENCE FROM N.A.
 RA Sverdlov S.D., Tsarev S.A., Markova S.V., Vasilenko S.K.,
 RA Chzhiyov V.E., Petrov N.A., Kusov Y.Y., Naestashenko T.A.,
 RA Balayan M.S.,
 RL Mol. Gen. Microbiol. Virol. 6:129-133(1987).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 DR EMBL: X15464; CAA33492.1; -
 DR InterPro: IPR000408; Reg. chr. condens.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 KW Polypeptide; Coat protein; Core protein; RNA-directed RNA polymerase;
 KW Hydrolase; Thiol protease.
 FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
 FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
 FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
 FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
 FT CHAIN 795 900 COAT PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 >1161 CORE PROTEIN P2C.
 FT NON TER 1161 1161
 SQ SEQUENCE 1161 AA; 131131 MW; 38B83789FEC3400 CRC64;

Query Match 100.0%; Score 990; DB 12; Length 1161;
 Best Local Similarity 100.0%; Pred. No. 4.1e-86;
 Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPYKEKRLLEVQKQRLKYAQSLSNEVLPKPKMKGLFSQAKISLFYEEHEIMK 60
 DB 786 SHIECKPYKEKRLLEVQKQRLKYAQSLSNEVLPKPKMKGLFSQAKISLFYEEHEIMK 845
 QY 61 FSWRGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKTEMKDKIVSLI 120
 DB 846 FSWRGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKTEMKDKIVSLI 905
 QY 121 EKFSTNKKYKSKVNFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMGL 180
 DB 906 EKFSTNKKYKSKVNFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMGL 965
 QY 181 SGVOEIKEQ 189
 DB 966 SGVOEIKEQ 974

RESULT 3
 ID Q9DLJ32 PRELIMINARY; PRT; 2225 AA.
 AC Q9DLJ32;
 DT 01-MAR-2001 (TREMblrel. 16, Created)

DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Polypeptide.
 OS Hepatitis A virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 CC NCB1_TaxID=12092;
 RN SEQUENCE FROM N.A.
 RC STRAIN=L-A-1;
 RA Wang P.F., Jiang C.L., Liu J.Y., Zhang H.Y.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF314208; AAG45423.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calici pol. hel.
 DR InterPro: IPR000408; Reg. chr. condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00910; RNA_helicase; 1.
 DR PRINTS: PR00918; CALICIVIRUSNS.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 SQ SEQUENCE 2225 AA; 251297 MW; EBACE41B043E5E9B CRC64;

Query Match 100.0%; Score 990; DB 12; Length 2225;
 Best Local Similarity 100.0%; Pred. No. 9.7e-86;
 Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPYKEKRLLEVQKQRLKYAQSLSNEVLPKPKMKGLFSQAKISLFYEEHEIMK 60
 DB 792 SHIECKPYKEKRLLEVQKQRLKYAQSLSNEVLPKPKMKGLFSQAKISLFYEEHEIMK 851
 QY 61 FSWRGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKTEMKDKIVSLI 120
 DB 852 FSWRGVTAADTRALRRFGFSLAAGRSVWTLMDAGVLTGRLIRLNDEKTEMKDKIVSLI 911
 QY 121 EKFSTNKKYKSKVNFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMGL 180
 DB 912 EKFSTNKKYKSKVNFPHGMLDLEIANSKDPFNMSETDLCFLHLWLNPKKINLADRMGL 971
 QY 181 SGVOEIKEQ 189
 DB 972 SGVOEIKEQ 980

RESULT 4

QY 09IFH5 PRELIMINARY; PRT; 2227 AA.
 AC Q9IFH5;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Polypeptide.
 OS Hepatitis A virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 CC NCB1_TaxID=12092;
 RN SEQUENCE FROM N.A.
 RC STRAIN=HAF-203;
 RA Baptista M.L., Silva M., de Lima M.A., Yoshida C.F., Gaspar A.M.,
 RA Pires Lopes M.O., Galler R.;
 RL "Nucleotide sequence of the HAF-203 hepatitis A virus strain isolated
 in Brazil and expression of the VP1 gene in a bacterial system.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF268396; AAF80114.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calici pol. hel.
 DR InterPro: IPR000408; Reg. chr. condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00910; RNA_helicase; 1.

DR PRINTS; PR00918; CALICIVIRUSNS.
DR POSITE; PS00626; RCCL_2; UNKNOWN 1.
SQ SEQUENCE 2227 AA; 251432 MW; 81913AEC8A04200 CRC64;

Query Match 99.7%; Score 987; DB 12; Length 2227;
Best Local Similarity 99.5%; Pred. No. 1.9e-85;
Matches 186; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVKQRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECRKPKYKELRLVGVKQRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTDADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDKXTKMDKIVSLI 120
DB 852 FSWRGVTDADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDKXTKMDKIVSLI 911
QY 121 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDPNNSETDLCFLHNLNPKKINLADRMIGL 180
DB 912 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDPNNSETDLCFLHNLNPKKINLADRMIGL 971
DB 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 5

Q9NMA2 PRELIMINARY; PRT; 2216 AA.
AC O9NMA2;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Polyprotein.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AH3;
RX MEDLINE=21386014; PubMed=11495028;
RA Fujitawa K., Yokosuka O., Fukai K., Imazeki F., Saisho H., Omata M.;
RT "Analysis of full-length hepatitis A virus genome in sera from
patients with fulminant and self-limited acute type A hepatitis.";
RL J. Hepatol. 35:112-119(2001).
DR EMBL: AB020566; BAA35104.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calici_pol_hel.
DR InterPro: IPR004008; Reg_chr_condens.
DR InterPro: IPR001205; RNA_helicase.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR POSITE; PS00626; RCCL_2; UNKNOWN 1.
SQ SEQUENCE 2216 AA; 250209 MW; 1A9D93FEC21FBE82 CRC64;

Query Match 99.4%; Score 984; DB 12; Length 2216;
Best Local Similarity 98.4%; Pred. No. 3.6e-85;
Matches 186; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVKQRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECRKPKYKELRLVGVKQRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTDADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDKXTKMDKIVSLI 120
DB 852 FSWRGVTDADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDKXTKMDKIVSLI 911
QY 121 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDPNNSETDLCFLHNLNPKKINLADRMIGL 180
DB 912 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDPNNSETDLCFLHNLNPKKINLADRMIGL 971

QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 6

ID Q67824 PRELIMINARY; PRT; 2218 AA.
AC Q67824;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE GBM/FRHK RNA.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GBM;
RX MEDLINE=94076453; PubMed=8254770;
RA Graff J., Normann A., Feinstein S.M., Flehmig B.;
RT "Nucleotide sequence of wild-type hepatitis A virus GBM in comparison
to two cell culture adapted variants.";
RL J. Virol. 68:548-554(1994).
DR EMBL: X75214; CA53024.1; -.
DR InterPro: IPR000408; Reg_chr_condens.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR POSITE; PS00626; RCCL_2; UNKNOWN 1.
FT CHAIN 1 785 P1 STRUCTURAL PROTEINS.
FT CHAIN 792 1422 P2 NONSTRUCTURAL PROTEINS.
FT CHAIN 1417 2218 P3 NONSTRUCTURAL PROTEINS.
SQ SEQUENCE 2218 AA; 250502 MW; CA72DF0922104C0E CRC64;

Query Match 99.4%; Score 984; DB 12; Length 2218;
Best Local Similarity 98.4%; Pred. No. 3.6e-85;
Matches 186; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVKQRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 786 SHIECRKPKYKELRLVGVKQRLKYAOEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 845
QY 61 FSWRGVTDADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDKXTKMDKIVSLI 120
DB 846 FSWRGVTDADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDKXTKMDKIVSLI 905
QY 121 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDPNNSETDLCFLHNLNPKKINLADRMIGL 180
DB 906 EKFTSNKYKSKVNFPHGMLDLEETIAANSKDPNNSETDLCFLHNLNPKKINLADRMIGL 965
QY 181 SGVOEIKEQ 189
DB 966 SGVOEIKEQ 974

RESULT 7

ID Q67825 PRELIMINARY; PRT; 2227 AA.
AC Q67825;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE GBM/WT RNA.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GBM;

RA MEDLINE=94076453; PubMed=8254770;
RA Grifft J., Normann A., Feinstone S.M., Flehmig B.;
RT "Nucleotide sequence of wild-type hepatitis A virus GBM in comparison
RT to two cell culture adapted variants.";
RL J. Virol. 68:548-554(1994).
DR EMBL; X75215; CAA53025.1; -
DR InterPro; IPR000408; Reg_chir_condens.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol.1.
DR Pfam; PF00910; RNA_helicase.1.
DR PROSITE; PS00626; RCCL_2; UNKNOWN.1.
FT CHAIN 1 791 P1 STRUCTURAL PROTEIN.
FT CHAIN 792 1422 P2 NONSTRUCTURAL PROTEIN.
FT CHAIN 1423 2227 P3 NONSTRUCTURAL PROTEIN.
SQ SEQUENCE 2227 AA; 251563 MW; 4C4D79D52F936B4 CRC64;

Query Match 99.4%; Score 984; DB 12; Length 2227;
Best Local Similarity 98.4%; Pred. No. 3.6e-85;

Matches 186; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVGKQRLKYAQEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECRKPKYKELRLVGVGKQRLKYAQEELSNEVLPPPRKMKGVFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTDTRALRRFGSILAAGRSVWTLMDAGVLTGRLRLINDEKWTMDDKIVSLI 120
DB 852 FSWRGVTDTRALRRFGSILAAGRSVWTLMDAGVLTGRLRLINDEKWTMDDKIVSLI 911
QY 121 EKFTSNKYKSKVNPFGHMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKYKSKVNPFGHMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 8

ID Q67826 PRELIMINARY; PRT; 2227 AA.
AC Q67826;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE GBM/HFS RNA.
OS Hepatitis A virus.
VR Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=12092;

RN [1] SEQUENCE FROM N.A.
RC STRAIN=GBM;
RX MEDLINE=94076453; PubMed=8254770;
RA Grifft J., Normann A., Feinstone S.M., Flehmig B.;
RT "Nucleotide sequence of wild-type hepatitis A virus GBM in comparison
RT to two cell culture adapted variants.";
RL J. Virol. 68:548-554(1994).
DR EMBL; X75216; CAA53026.1; -
DR InterPro; IPR000408; Reg_chir_condens.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol.1.
DR Pfam; PF00910; RNA_helicase.1.
DR PROSITE; PS00626; RCCL_2; UNKNOWN.1.
FT CHAIN 1 791 P1 STRUCTURAL PROTEIN.
FT CHAIN 792 1422 P2 NONSTRUCTURAL PROTEIN.
FT CHAIN 1423 2227 P3 NONSTRUCTURAL PROTEIN.
SQ SEQUENCE 2227 AA; 251496 MW; 488CB7C962319457 CRC64;

Query Match 99.4%; Score 984; DB 12; Length 2227;
Best Local Similarity 98.4%; Pred. No. 3.6e-85;
Matches 186; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVGKQRLKYAQEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECRKPKYKELRLVGVGKQRLKYAQEELSNEVLPPPRKMKGVFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTDTRALRRFGSILAAGRSVWTLMDAGVLTGRLRLINDEKWTMDDKIVSLI 120
DB 852 FSWRGVTDTRALRRFGSILAAGRSVWTLMDAGVLTGRLRLINDEKWTMDDKIVSLI 911
QY 121 EKFTSNKYKSKVNPFGHMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKYKSKVNPFGHMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 9

ID Q9WMA3 PRELIMINARY; PRT; 2227 AA.
AC Q9WMA3;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polypotein.
OS Hepatitis A virus.
VR Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=12092;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=AH2;
RX MEDLINE=21386014; PubMed=11495028;
RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saito H., Omata M.;
RT "Analysis of full-length hepatitis A virus genome in sera from
RT patients with fulminant and self-limited acute type A hepatitis.";
RL J. Hepatol. 35:112-119(2001).
DR EMBL; AB020565; BAA35103.1; -
DR MEROPS; C03.005; -
DR InterPro; IPR004004; Galici_pol_hel.
DR InterPro; IPR000408; Reg_chir_condens.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol.1.
DR Pfam; PF00910; RNA_helicase.1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR PROSITE; PS00626; RCCL_2; UNKNOWN.1.
SQ SEQUENCE 2227 AA; 251440 MW; E04F846CEC7F50FD CRC64;

Query Match 99.4%; Score 984; DB 12; Length 2227;
Best Local Similarity 98.4%; Pred. No. 3.6e-85;
Matches 186; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLVGVGKQRLKYAQEELSNEVLPPPRKMKGLFSQAKISLFYTEHEIMK 60
DB 792 SHIECRKPKYKELRLVGVGKQRLKYAQEELSNEVLPPPRKMKGVFSQAKISLFYTEHEIMK 851
QY 61 FSWRGVTDTRALRRFGSILAAGRSVWTLMDAGVLTGRLRLINDEKWTMDDKIVSLI 120
DB 852 FSWRGVTDTRALRRFGSILAAGRSVWTLMDAGVLTGRLRLINDEKWTMDDKIVSLI 911
QY 121 EKFTSNKYKSKVNPFGHMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKYKSKVNPFGHMLDEEIAANSKDFPNMSETDLCFLHMLNPKKINLADRMIGL 971
QY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 10

Q9WMAO

ID O9WMA0 PRELIMINARY; PRT; 2227 AA.
AC O9WMA0
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polypeptide.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FM3;
RX MEDLINE=21386014; PubMed=11495028;
RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saisho H., Omata M.;
RT "Analysis of full-length hepatitis A virus genome in sera from
patients with fulminant and self-limited acute type A hepatitis.";
RL J. Hepatol. 35:112-119(2001).
DR EMBL; AB020568; BAA35106.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR004008; Reg_chir_condens.
DR InterPro; IPR001205; RNA_helicase.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
DR PROSITE; PS00626; RCCL_2; UNKNOWN 1.
SQ SEQUENCE 2227 AA; 25118 MW; 53B6B4432127B9B CRC64;

Query Match 99.1%; Score 981; DB 12; Length 2227;
Best Local Similarity 98.4%; Pred. No. 7.1e-85;
Matches 186; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGKORLKYAOEELSNVLPKPKKGLFSQAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKELRLEVGKORLKYAOEELSNVLPKPKKGLFSQAKISLFTTEHEIMK 851
OY 61 FSWRGVADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDEKWTMCKDIXVSLI 120
DB 852 FSWRGVADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDEKWTMCKDIXVSLI 911
OY 121 EKFTSNKTKWTKVNPFGMLDLEETIAANSKDPNNSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKTKWTKVNPFGMLDLEETIAANSKDPNNSETDLCFLHMLNPKKINLADRMIGL 971
OY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980
RESULT 11
O9WMA0 PRELIMINARY; PRT; 2227 AA.
AC O9WMA0
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polypeptide.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FM3;
RX MEDLINE=21386014; PubMed=11495028;
RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saisho H., Omata M.;
RT "Analysis of full-length hepatitis A virus genome in sera from
patients with fulminant and self-limited acute type A hepatitis.";
RL J. Hepatol. 35:112-119(2001).
DR EMBL; AB020568; BAA35107.1; -.
DR MEROPS; C03.005; -.

DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR004008; Reg_chir_condens.
DR InterPro; IPR006005; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
DR PROSITE; PS00626; RCCL_2; UNKNOWN 1.
SQ SEQUENCE 2227 AA; 251387 MW; 817640187672A23D CRC64;

Query Match 98.8%; Score 978; DB 12; Length 2227;
Best Local Similarity 98.4%; Pred. No. 1.4e-84;
Matches 186; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGKORLKYAOEELSNVLPKPKKGLFSQAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKELRLEVGKORLKYAOEELSNVLPKPKKGLFSQAKISLFTTEHEIMK 851
OY 61 FSWRGVADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDEKWTMCKDIXVSLI 120
DB 852 FSWRGVADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDEKWTMCKDIXVSLI 911
OY 121 EKFTSNKTKWTKVNPFGMLDLEETIAANSKDPNNSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKTKWTKVNPFGMLDLEETIAANSKDPNNSETDLCFLHMLNPKKINLADRMIGL 971
OY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 12
O8OV03 PRELIMINARY; PRT; 2227 AA.
ID O8OV03
AC O8OV03
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polypeptide.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LU38/WT;
RA Hu Y., Hu N.;
RT "Hepatitis A Virus LU38/WT";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF357222; AAM08224.1; -.
SQ SEQUENCE 2227 AA; 251379 MW; 975C48511E9213D7 CRC64;

Query Match 98.8%; Score 978; DB 12; Length 2227;
Best Local Similarity 97.9%; Pred. No. 1.4e-84;
Matches 185; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGKORLKYAOEELSNVLPKPKKGLFSQAKISLFTTEHEIMK 60
DB 792 SHIECRKPKYKELRLEVGKORLKYAOEELSNVLPKPKKGLFSQAKISLFTTEHEIMK 851
OY 61 FSWRGVADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDEKWTMCKDIXVSLI 120
DB 852 FSWRGVADTALRRFGFSLAAGRSVWTLNDAGVLTGRLIRLNDEKWTMCKDIXVSLI 911
OY 121 EKFTSNKTKWTKVNPFGMLDLEETIAANSKDPNNSETDLCFLHMLNPKKINLADRMIGL 180
DB 912 EKFTSNKTKWTKVNPFGMLDLEETIAANSKDPNNSETDLCFLHMLNPKKINLADRMIGL 971
OY 181 SGVOEIKEQ 189
DB 972 SGVOEIKEQ 980

RESULT 13
 Q9WMA1 PRELIMINARY; PRT; 2227 AA.
 AC Q9WMA1; 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Polypeptide.
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F.H.
 RX MEDLINE=21386014; PubMed=11495028;
 RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saisho H., Omata M.;
 RT "Analysis of full-length hepatitis A virus genome in sera from
 RT patients with fulminant and self-limited acute type A hepatitis.";
 J. Hepatol. 35:112-119(2001).
 EMBL: AB020567; BAA35105.1; -.
 DR MEROPS: C03.005; -.
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR000408; Reg_chir_condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00910; RNA_helicase_1.
 DR PRINTS: PR00918; CALICIVIRUSNS.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 SQ SEQUENCE 2227 AA; 251415 MW; F92C8E2323FC5621 CRC64;

Query Match 98.7%; Score 977; DB 12; Length 2227;
 Best Local Similarity 98.4%; Pred. No. 1.7e-84;
 Matches 186; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SHIECRKPYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKLFSGQAKISLFYTEHEIMK 60
 DB 792 SHIECRKPYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKLFSGQAKISLFYTEHEIMK 851
 QY 61 FSWRGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKTEKDDKIVSLI 120
 DB 852 FSWRGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKTEKDDKIVSLI 911
 QY 121 EKFTSNKYKSKVPFGHMLDEEIAANSKDPNNSETDLCFLHWNPKKINLADRMGL 180
 DB 912 EKFTSNKYKSKVPFGHMLDEEIAANSKDPNNSETDLCFLHWNPKKINLADRMGL 971
 DB 181 SGVQEIKEQ 189
 DB 972 SGVQEIKEQ 980

RESULT 14
 Q67817 PRELIMINARY; PRT; 2218 AA.
 AC Q67817; 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Complete genome.
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F.G.
 RX MEDLINE=95381623; PubMed=7653108;
 RA Benenue F., Pisani G., Divizia M., Pana A., Morace G.;
 RT "Complete nucleotide sequence of a cytopathic hepatitis A virus strain
 RT isolated in Italy.";
 Virus Res. 36:299-309(1995).
 RL

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F.G.
 RA Morace G.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X83502; CA58281.1; -.
 DR InterPro: IPR000408; Reg_chir_condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00910; RNA_helicase_1.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 SQ SEQUENCE 2218 AA; 250476 MW; 813B21D3E4E533CA CRC64;

Query Match 98.5%; Score 975; DB 12; Length 2218;
 Best Local Similarity 97.9%; Pred. No. 2.6e-84;
 Matches 185; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHIECRKPYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKLFSGQAKISLFYTEHEIMK 60
 DB 766 SHIECRKPYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKLFSGQAKISLFYTEHEIMK 845
 QY 61 FSWRGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKTEKDDKIVSLI 120
 DB 846 FSWRGVYADTRALRRFGSLAAGRSVWTLMDAGVLTGRLRLNDEKTEKDDKIVSLI 905
 QY 121 EKFTSNKYKSKVPFGHMLDEEIAANSKDPNNSETDLCFLHWNPKKINLADRMGL 180
 DB 906 EKFTSNKYKSKVPFGHMLDEEIAANSKDPNNSETDLCFLHWNPKKINLADRMGL 965
 QY 181 SGVQEIKEQ 189
 DB 966 SGVQEIKEQ 974

RESULT 15
 Q9WMA4 PRELIMINARY; PRT; 2227 AA.
 AC Q9WMA4; 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Polypeptide.
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AH1;
 RX MEDLINE=21386014; PubMed=11495028;
 RA Fujiwara K., Yokosuka O., Fukai K., Imazeki F., Saisho H., Omata M.;
 RT "Analysis of full-length hepatitis A virus genome in sera from
 RT patients with fulminant and self-limited acute type A hepatitis.";
 J. Hepatol. 35:112-119(2001).
 EMBL: AB020564; BAA35102.1; -.
 DR MEROPS: C03.005; -.
 DR InterPro: IPR004004; Calici_pol_hel.
 DR InterPro: IPR000408; Reg_chir_condens.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00910; RNA_helicase_1.
 DR PRINTS: PR00918; CALICIVIRUSNS.
 DR PROSITE: PS00626; RCCL_2; UNKNOWN_1.
 SQ SEQUENCE 2227 AA; 251304 MW; 0DEF6D2AEC29C0CE CRC64;

Query Match 98.3%; Score 973; DB 12; Length 2227;
 Best Local Similarity 97.9%; Pred. No. 4.1e-84;
 Matches 185; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SHIECRKPYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKLFSGQAKISLFYTEHEIMK 60
 DB 766 SHIECRKPYKELRLEVGKQRLKYAOEELSNEVLPPPRKMGKLFSGQAKISLFYTEHEIMK 845

Db	792	SHIESRKPYKEIRLEVGKQRLKYAQEBLSNEVLPPPRKIKGLFSQAKISLFYTEHEIMK	851
Qy	61	FSWRCVTADTRALRRFRFGSLAAGRSVWTLMDAGVLTGRLIRLNDEKWTMCKDKI VSLI	120
Db	852	FSWRCVTADTRALRRFRFGSMAGRSVWTLMDAGVLTGRLVRLNDEKWTMCKDKI VSLI	911
Qy	121	EKFTSNKYWSKYNPPHMLDLEIANSKDFPNMSETDLCFLHMLNPKKINLADRMGL	180
Db	912	EKFTSNKYWSKYNPPHMLDLEIANSKDFPNMSETDLCFLHMLNPKKINLADRMGL	971
Qy	181	SGVOEIXEQ	189
Db	972	SGVOEIXEQ	980

Search completed: April 2, 2003, 11:10:31
Job time : 56.7212 secs

THIS PAGE BLANK (USPTO)